Arabic in Contact

Edited by Stefano Manfredi
and Mauro Tosco

6

Arabic in Contact

Studies in Arabic Linguistics

This book series aims to publish original research in all fields of Arabic linguistics, including – but not limited to – theoretical linguistics, historical linguistics, sociolinguistics, pragmatics, typology, and language acquisition. Submissions from all current theoretical frameworks are welcome. Studies may deal with one or more varieties of Arabic, or Arabic in relation to or compared with other languages. Both monographs and thematic collections of research papers will be considered.

The series includes monographs and thematically coherent collective volumes, in English.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/sal

Editors

Volume 6

Arabic in Contact
Edited by Stefano Manfredi and Mauro Tosco

Arabic in Contact

Edited by

Stefano Manfredi
CNRS, SeDyl

Mauro Tosco
University of Turin

Table of contents

Section 3.   Deep contact: Arabic-based contact languages

Section 4.   Back to the speaker: Codeswitching and language ideologies

Arabic in contact, now and then

Stefano Manfredi and Mauro Tosco
SeDyL, CNRS / University of Turin

1. Contact linguistics and Arabic in contact[1]

Of course, languages are *not* in contact. We could say that *speakers* of languages are, but even this would be misleading, as we would bestow an undue role to individuals in their capacity as speakers, and at the same time forget that words and patterns spread when they are heard, rather than when they are uttered. As Thomason (2001: p. 2) puts it, "in the simplest definition, language contact is the use of more than one language in the same place at the same time". Thus, what we call language contact is one facet of human interaction, never separable from it. An obvious factor favoring language contact is widespread bi- and multilingualism. However, chances of language contact might be increased/decreased among other things, by the relative number of speakers of a given language, their geographical location, their movement opportunities, and the technologies they use in communication – all the way down to the individual disposition to contact. Still, in any case, individuals interact, either face to face or not. In this respect, it is also important to remark that when we call an individual "a speaker of [language] X" we qualify them on the basis of their verbal behavior only – leading us eventually to forget that language interaction is ultimately just one of the results of human interaction (verbal and non-verbal alike).

Many individuals have been in contact with speakers of Arabic for a long time, in different parts of the world and for different reasons, more than we can explore, or even mention, here. One of the reasons for contact with Arabic all around the world is linked to the role of Classical Arabic as the language of *Islām*, especially in Africa and Asia. This cultural aspect of the spread of Arabic produces an indirect type of language contact that will not be explored here (see Versteegh 2015; Tosco 2015). Equally, largely absent from our survey will be the issues of dialect contact (i.e., dialect levelling and dialect mixing, see Miller et al. 2007), diglossic bilingualism involving Modern Standard Arabic and language contact in diasporic contexts (see Rouchdy 2002). A further dimension of language contact which will not be covered in the present volume is the role of Arabic as a source of neologisms in language planning: in cases such as contemporary Ethiopia, Arabic words may be preferred over older loans from a local dominant language (such as Amharic) and purposefully imposed in the new, standard written languages (e.g., in Oromo; Savà and Tosco 2008). In this case, actual contact between speakers is largely immaterial.

More humbly, the aim of the present volume is to provide an overview of current trends in the study of language contact *involving* Arabic. If 'contact' refers to contact between speakers, to separately investigate its effects as unfolding *in* and *from* Arabic is largely artificial. By drawing on the social factors that have converged to create different contact situations, we therefore concentrate on both contact-induced change in Arabic and language change through contact with Arabic. Furthermore, we aim at covering other important aspects related to language contact involving Arabic, such as the emergence of Arabic-based contact varieties, codeswitching, and metalinguistic representations of contact-induced changes.

For a long time, scholars have tried to typologize the outputs of language contact in light of both different contact situations and the nature of the linguistic structures in contact. In this regard, Weinreich (1953: p. 86) overtly states that the ultimate goal of contact linguistics is "to predict typical forms of interference from the sociolinguistic description of a bilingual community and a structural description of its languages". The most influential theoretical paradigm on language contact has probably been suggested by Thomason and Kaufman (1988) who provide for three main contact scenarios. The first one is that of language maintenance which typically implies "borrowing" or, in other words, the incorporation of foreign elements into the speaker's native language. The second scenario, that of language shift, is related to (substrate) "interference", which is instead conceived as the linguistic influence played by an ancestral language over an intrusive language that gradually supplants it. The third scenario involves the creation of new linguistic systems composed of elements of different languages in contact and it corresponds to pidginization/creolization as well as to language mixing. The idea underlying this sociohistorical understanding of language contact is that contact-induced change

can occur at any level of a given linguistic system. However, the processes of "borrowing" and "interference" differ sharply in terms of linguistic outputs, the former being mainly related with the transfer of lexical and morphological material, and the latter inducing the transfer of phonological, syntactic and semantic constraints.

Several studies of language contact involving Arabic (Thomason 2006; Versteegh 2001, 2010) have been primarily inspired by the explanatory model of Thomason and Kaufman. Despite this, the traditional sociohistorical understanding of contact-induced change has also been criticized. For instance, Myers-Scotton (2002) argues that different contact phenomena result from a limited set of grammatical processes, regardless of the sociolinguistic scenario in which they take place. Lucas (2012: p. 521), on his part, aptly observes that it is far from clear that the question of whether or not a community happens to maintain its ancestral language is crucial to understanding the dynamics of contact-induced change. As a further matter, it has been repeatedly observed that there is no clear-cut line between borrowing and interference (Haspelmath 2009). Aikhenvald (2007: p. 4), for example, defines "borrowing" as "the transfer of features of any kind from one language to another as the result of contact". and "interference" as "the non-deliberate carrying over of linguistic features from one's first language into one's second language". In this acceptation, "interference" is nothing more than a subtype of "borrowing".

In the light of the above, over the last few decades scholars have advocated viewing the outcomes of language contact from other perspectives. Typological research in language contact focuses on the interplay of two or more linguistic systems in order to compare the effects of contact on language structures (Matras 2001). Adopting this typological standpoint, Ross (2006, 2007) detaches himself from the tripartite conception of contact-induced change proposed by Thomason and Kaufman and eventually distinguishes between two main processes of language transfer: "typical borrowing" and "typical shift-induced interference". On the one hand, the process of typical borrowing is produced by native speakers who intentionally import lexical items from another language into their own language. On the other hand, typical shift-induced interference is produced by bilingual speakers who unconsciously import lexical and grammatical features of a dominant language into their own ancestral language. In such situations, bilingual speakers tend to transfer syntactic constructions from the socially dominant language, resulting in a contact-induced typological change labelled as "metatypy". Metatypy often presupposes a high degree of bi- and multilingualism among the members of a group, with the ancestral language being the intragroup means of communication, and the socially dominant language being used for intergroup communication. As far as the study of Arabic in contact is concerned, the typological notion of metatypy has been adopted for describing the contact-induced typological change affecting both minority varieties of Arabic, such as the Central-Asian dialect of Bukhara

(Ratcliffe 2005), and minority languages in contact with Arabic, as in the case of Laggorí, an Eastern-Sudanic (Daju) language spoken in the Nuba Mountains region in Sudan (Manfredi 2014).

A third prominent theoretical framework for the study language contact is that proposed by Van Coetsem (1988, 1995) and further developed by Winford (2005, 2007). Unlike previous approaches, Van Coetsem's explanatory model of language contact is neither socio-historically nor typologically oriented, since it rather focuses on the psycholinguistic criterion of "language dominance" (see also Smits 1998). According to Van Coetsem, a bilingual speaker is dominant in the language in which they are most proficient and that is not necessarily their native language or the socially dominant language. Against this backdrop, he proposes two distinct transfer types: "borrowing", which is typically produced by speakers who are dominant in the recipient language, and "imposition", which is instead produced by speakers who are dominant in the source language (corresponding to Thomason and Kaufman's concept of donor language). Moreover, Van Coetsem (1988: p. 20; 1995: p. 25) and Winford (2005: p. 377) point out that the dissimilar outcomes of borrowing and imposition are primarily a result of the "stability gradient" of language, which induces speakers to preserve the domains of their dominant language that are less affected by change. This is the main reason borrowing tends to be irregular and typically involves the transfer of lexical items, whereas imposition is more systematic and produces significant grammatical changes. Despite this, it is not always a trivial matter to tease the two transfer types apart since bilingual speakers may trigger borrowing and imposition in the same contact situation while directing them towards different languages. Crucially, in contrast to the traditional sociohistorical standpoint represented by Thomason and Kaufman, Winford (2005: p. 396; 2008: p. 128) assumes that the processes that create contact languages are the same as those that operate in contact-induced change. Given this background, three broad categories of contact languages may be identified: contact languages that primarily arose through borrowing (such as the case of Maltese), languages that primarily arose through imposition (such as the case Arabic-based creoles), and languages that arose from a combination of both transfer types (e.g., Central-Asian Arabic). Only a small number of comprehensive studies have hitherto adopted Van Coetsem's psycholinguistic model of language contact to Arabic (Lucas 2012, 2014; Manfredi 2018) and to the influence of Arabic on other languages (Kossmann 2013b).

The review of the main frameworks of language contact presented above is far from being exhaustive. Still, it gives an idea of the multiplicity of theoretical standpoints on language contact and their respective impacts on the study of the dynamic of language contact involving Arabic. In this light, we believe that regardless of the approach one adopts, language contact is above all a multifactorial process

of language change (Chamoreau and Léglise 2012), encompassing, inter alia, sociolinguistic, typological and psycholinguistic factors. For the sake of the present volume, we do not align ourselves with any preferred model of contact-induced change, as we prefer to leave contributors free to adopt the most suitable approach for their own studies. The following sections detail the rationale of the contents of this volume.

2. In and from Arabic: Grammar in context

Grammatical borrowing involves the transfer of grammatical structures from a donor language to a recipient language. It is now widely agreed that grammatical borrowing entails the transfer of a wide range of segmental grammatical structures (e.g. free and bound morphemes) and non-segmental ones (e.g. syntactic and semantic constraints). The comparative study of grammatical borrowing must therefore take into account both the "horizontal" diversity of languages in contact, and the "vertical" diversity of the grammatical categories on which contact can have an impact (Matras and Sakel 2007: p. 2). In the case of Arabic, the outputs of grammatical borrowing have been traditionally analyzed in terms of the substratal interference with modern Arabic dialects on the part of, inter alia, Himyarite (Diem 1979), Aramaic (in its different varieties, see for example Contini 1999), Coptic (Lucas and Lash 2010), and Berber (again, in different forms and times, see for example Taine-Cheikh 2008; see Kossmann 2013a for a general overview). A smaller number of studies focus on the grammatical influence of Arabic on other languages (see for example Arnold 2007; Matras 2007; Kossmann 2013b; Souag 2014; Coghill 2015). In Section 1, we concentrate on the grammatical effects of contact, in cases where it induces changes in Arabic, as well as where contact with Arabic induces change elsewhere.

In 'The Arabic component in Domari', Bruno Herin investigates the poorly-documented (and by now largely displaced and severely endangered) Domari, an Indo-Aryan language spoken by Dom people, on the basis of his own largely unpublished fieldwork and supplementing Matras' (2012) extensive investigation of Palestinian Domari. Domari is a primary example of language contact, insofar as its speakers are and have traditionally been bilingual. Arabic is just the most recent among a large number of languages with which Domari has interacted, with very different results: generally, Herin shows that Arabic influence on Domari has been stronger in the south (Palestine) than in the north (Syria, Lebanon and southern Turkey). Such a differential impact (leading in extreme cases, as in Jerusalem, to language shift to Arabic by a majority of Domari speakers) is visible in morphology, syntax and lexicon, and can be characterized as leading mainly to pattern

replication in the north and matter replication in the south. This in turn seems to suggest, in the author's words, that "phenomena such as bilingual suppletion in particular and large scale transfer of matter in general involve a greater historical depth of bilingualism and a more advanced stage of language attrition" in the south.

It is instead an Arabic variety which is the target of contact in Faruk Akkuş and Elabbas Benmamoun's 'Syntactic outcomes of contact in Sason Arabic (Turkey)'. An endangered variety of southern Turkey, Sason Arabic has been in contact with both Turkish and (possibly for a much longer period) Kurdish. The authors concentrate on indefiniteness, light verb constructions, causatives, and negative copula sentences, and show how Sason Arabic patterns with the languages it is in contact with rather than with Arabic at large, making contact as the most plausible source of pattern change.

We move to Africa and to a very different contact pattern with Lameen Souag's 'Arabic-Berber-Songhay contact and the grammaticalisation of 'thing''. Souag investigates the development of double negation in Arabic, focusing once again on the striking parallels between Berber and North African Arabic in this domain. The crux of the matter revolves around the contact-induced grammaticalization (Heine and Kuteva 2003; 2005) of reflexes of (Classical) Arabic *šayʔ* as a negation marker, but also in indefinite quantification and polar question marking, both across North African dialects of Arabic and Berber languages. The author proposes a relative chronology of these developments and points out how non-Arabic varieties sometimes preserve usages which are obsolete in present-day Arabic dialects.

We remain within the dynamics of Berber–Arabic contact with Dominique Caubet's contribution 'Arabic and Berber in contact: Arabic in a minority situation in El Hoceima Region'. The author introduces us to a complex and somehow paradoxical contact situation in an area of northern Morocco, where both Jebli Moroccan Arabic and Tarifit Berber are spoken. Thus, within one single faction of one single tribe, we find both Arabic and Berber speakers; in particular, we have a *minority* of Arabic speakers among the Berber-speaking Aït Aïssa faction, who are themselves a *minority* within the mostly Arabic-speaking Beni Iṭṭeft tribe. The situation seems to be one of stable bilingualism, going on with little changes since it was first studied in 1932, and possibly for centuries before that.

In 'Arabic on the Dahlak islands (Eritrea)' Marie-Claude Simeone-Senelle describes a very poorly documented Arabic variety of the Red Sea based on her unpublished material, and compares it to the Arabic spoken along the African coast as a *lingua franca* (Simeone-Senelle 1999). Generally speaking, a certain amount of morphological reduction occurs in non-native varieties of Arabic. However, vehicular varieties of Arabic in Africa present a lower degree of grammatical restructuring in comparison with Arabic-based pidgins and creoles (Manfredi 2013; Tosco and Manfredi 2013, *cf.* 4). Trying to determine to what extent a distinction can be

drawn between native and non-native varieties of Dahlak Arabic, Simeone-Senelle's results show that the vernacular and vehicular forms of Arabic tend to merge and level into a single local variety. In this context, the role played by Dahalik, the Ethio-Semitic language dominant on the islands, is not as significant as it could be expected and appears to be dwindling.

3.   In and from Arabic: Dealing with words

Loanwords are the most obvious result of language contact. This is simply because, being highly referential, lexical categories are more likely to be borrowed than grammatical categories. Following Haspelmath (2009: p. 36), we define "loanword" as a lexical item that at some point in the history of a language entered its lexicon as a result of "borrowing", here intended as an umbrella term for all kinds of transfer from a donor language to a recipient language. The probability of lexical borrowing depends, of course, on both linguistic and extra-linguistic factors. On the one hand, high token frequency may play an important role in triggering the integration of a given lexical item. On the other, negative language attitudes and linguistic purism can limit, or even hinder, lexical borrowing. A large body of literature has been devoted to lexical borrowing involving Arabic (see Versteegh 2001; 2010 for a general overview) and a good number of contributions in this volume deal, one way or the other, with loanwords – again, both in and out of Arabic, and both "now" and "then".

These dimensions are tackled in Section 2, starting with Catherine Taine-Cheikh's 'Ḥassāniyya Arabic in contact with Berber: the case of quadriliteral verbs'. Ḥassāniyya Arabic is the dominant language in Mauritania, while Berber (specifically, the Zenaga Berber language) is highly endangered. The author's analysis demonstrates that Ḥassāniyya Arabic has incorporated a good deal of four-consonant roots of Berber origin. Even more common is the case of new formations from Berber nominal borrowings. These pertain to the category of "cultural borrowings" rather than "core borrowings", with two semantic fields dominating: animals (husbandry, riding, doctoring), and illnesses, followed by traditional activities, physical traits, social features and time-related vocabulary (*cf.* Myers-Scotton 2002: p. 41; see Haspelmath 2009: p. 46–50 for a critical review of "core" and "cultural" borrowings). While Ḥassāniyya often retains Berber loanwords which are absent or disappeared in Zenaga itself, various semantic shifts have made their appearance, often together with semantic specialization.

A typical example of lexical borrowing in Arabic as a native recipient language is represented by loanwords from European languages into modern Arabic dialects. In 'Loan verbs in Egyptian Arabic: new findings and evidence from social media', Ashraf Hassan explores the increasing lexical impact of English on Egyptian

Arabic focusing on the morphophonological integration of loan verbs in the context of social media. Drawing on the typology of verbal borrowings proposed by Wohlgemuth (2009), Hassan explores the different accommodation mechanisms of loan verbs in written Egyptian Arabic (light verb strategy vs. direct insertion) and proposes a diachronic explanation for their variable incidence.

Luca D'Anna, in 'Italian loanwords in Libyan Arabic: morphophonological analysis and semantic considerations', revisits and updates, building at least partially upon unpublished fieldwork, the long history of contact between Italian and Libyan Arabic. Contact has reached its apex before and during colonial times in the first half of the 20th century (Italy conquered Libya in 1911 and a substantial Italian colony was present in the country until 1969). There are currently at least 700 lexical items in Libyan Arabic that can be traced back to borrowings from Italian (but for some of them the source may also be some other Romance language). Based on quantitative and qualitative observations, the paper goes through both the phono-morphological and semantic integration of these borrowed items.

As is well known, Arabic represents an important source of loanwords in African languages. In this vein, Nicolas Quint proposes 'An assessment of the Arabic lexical contribution to contemporary spoken Koalib (Sudan)', a Niger-Kordofanian language spoken by approximately 100,000 people in the Nuba Mountains, in western Sudan. Contact with Arabic is at least 250 years old and commenced with the arrival of Arabic-speaking nomads and the rise of a local Muslim kingdom. All varieties of Koalib, even the most conservative, have witnessed the integration of a sizable number of Arabic items belonging to different parts of speech. A corpus of approximately 300 Koalib items borrowed from different Arabic varieties (mainly from Kordofanian Baggara Arabic and Sudanese Colloquial Arabic) is thoroughly analyzed as far as the phonology, morphology and semantics are concerned. Similar to Beja (Vanhove 2012), a northern Cushitic language spoken in eastern Sudan, the integration of Arabic lexical items in Koalib may also entail the copying of productive morphological patterns of the donor language. This shows how contact-induced morphological innovations are usually transferred into the recipient language via lexical borrowing (King 2000).

4.   Deep contact: Arabic-based contact languages

Apart from the aforementioned types of contact-induced change, Arabic has been involved in the emergence of a number of contact languages. According to Bakker and Matras (2013: p. 1), the notion of "contact languages" generally refers back to "new languages that have emerged in extreme contact situations where available language repertoires did not provide an effective tool for communication". Broadly

speaking, scholars identify three types of contact languages: pidgins, creoles and mixed languages. As far as pidgin and creole languages are concerned, despite the unfortunately recalcitrant belief that they merely represent simplified versions of their lexifier languages, the only valid criterion for defining them against languages arising out of "normal" language transmission lies very possibly only in the social conditions for their emergence. As a matter of fact, pidgins and creoles are different from other spoken languages in that they came into existence as a consequence of the disruption of the intergenerational transmission of the lexifier language (Comrie 2011: p. 600). These uncommon conditions of language emergence entail different processes of language change linked to second language acquisition with limited input, substratum interference, as well as to internal developments (Winford 2005: p. 411). Mixed languages, on their part, are conventionally seen as products of extensive bilingualism whose grammar and lexical systems can be traced back to more than a single source language (Matras and Bakker 2003).[2] It is for this very reason that the source of pidgin/creole language structures are generally opaque, whereas those of mixed languages are relatively transparent (Owens 2001: p. 53).

Section 3 is largely, but not exclusively, concerned with Arabic-based pidgins and creoles. At the same time, Arabic language mixing, such as found in Maltese and Central Asian Arabic, is not considered. Attention on Arabic-based pidgins and creoles (and what lies between) has received a good deal of attention in recent years. The editors of this volume had previously published a collection of articles addressed at creolists and general linguists (Manfredi and Tosco 2014) and an overview for scholars in Arabic (Tosco and Manfredi 2013). Against the backdrop of this growing amount of data on Arabic contact languages, in the present volume we prefer to concentrate on theoretical issues with contributions by Jonathan Owens and Kees Versteegh, followed by two data-oriented articles by Andrei Avram and Shuichiro Nakao.

The question 'Why linguistics needs a historically oriented Arabic linguistics' is addressed by Jonathan Owens. The author applies Labov's (2007) distinction between transmission and diffusion (while the former results in gradual incremental changes, the latter yields larger and irregular change), to the study of five Arabic cases: Emirati, Nigerian, Baghdadi, Uzbekistan (Central Asian) Mixed Arabic, and Nubi. Arabic, like American English in Labov's study, shows striking language stability across geo-diachronically widely separated varieties, as well as impressive cases of widespread contact-induced change, but these can be considered irregular only in the case of Nubi (a creole). The study highlights how global criteria for

---

2. In this sense, every language could be considered to be "mixed" in that it presents some lexical or grammatical element deriving from another language. As a consequence, some scholars openly argue against the operativeness of the notion of "mixed language" (Versteegh 2017).

defining the outcomes of transmission vs. diffusion remain elusive, and most of all, offers interesting insights into the workings of historical linguistic processes offered by Arabic and its rich and variegated history.

In 'Basic varieties of Arabic', Kees Versteegh applies the model of the Basic Variety (developed by Klein and Perdue 1997 and further elaborated by Benazzo 2003) to two basic forms of communication in Arabic, Pidgin Madame and Gulf Pidgin Arabic. Predictions on the sequentiality of development of temporal adverbs of contrast (resultative *already*; continuative *still*) based upon Benazzo's analysis of the Basic Varieties of German, French and English fail to be supported when Arabic is taken into consideration: although the source language of these two Arabic Basic Varieties does not contain a resultative adverb, both varieties feature it as *kalas*. Both this and the relatively frequent use of a continuative particle (*bād*) at a very early stage contradict the universality of Benazzo's results.

Not much is yet known about Arabic foreigner talk and its role in the emergence of Arabic-based pidgins in the Middle-East. In 'On the relationship between Arabic Foreigner Talk and Pidgin Arabic', Andrei Avram compares the morphosyntax and lexical features of the Arabic Foreigner Talk to those of four Arabic-lexifier pidgins (Pidgin Madame, Jordanian Pidgin Arabic, Romanian Pidgin Arabic, and Gulf Pidgin Arabic). The author proposes a feedback relationship in order to account for the significant number of features shared between Arabic Foreigner Talk shares and all or at least some of these Arabic-based pidgins.

In 'A 'creole' music in Juba Arabic: *direr* dance in Juba (South Sudan)', Shuichiro Nakao challenges the assumption that Nubi and Juba Arabic, the two Arabic creoles spoken in Eastern Africa, have been cut off from each other since their early divergence in the 1880s, when Anglo-Sudanese troops stationed in modern-day Southern Sudan were forced to move south and settle in Uganda and Kenya in the wake of the Mahdist revolution in Sudan. Nakao presents ethnographic evidence to the contrary, and, as a foremost example of these "inter-creole" contacts, explores the musical practice called *dolúka* in Nubi and *dirêr* in Juba Arabic. The results show that across eastern Africa, Arabic-based speech communities have been very much in contact through most of their history, shaping and continually redefining their identity through language and culture contacts.

5.   Back to the speaker: Codeswitching and language ideologies

The last section of the volume deals with the speaker's involvement and processing in language contact. The focus here is on speaker role as both producer and conceptualizer of contact-induced change. The first two contributions deal with codeswitching involving Arabic. Codeswitching is not a kind of diachronic

contact-induced language change, but rather a type of synchronic contact-induced speech behavior (Hasplemath 2009: p. 40) which may have long-run implications in terms of linguistic convergence (Muysken 2000). In this perspective, codeswitching clearly differs from the integration of lexical and grammatical loans. In spite of this, during the last decades the traditional notion of codeswitching has been facing a growing criticism. Clyne (2003: p. 72) affirms that "the term 'codeswitching' has now become so polysemous and unclear that it is necessary to find more precise terms to map out the boundaries and interfaces". Along the same lines, Winford (2003: pp. 107–108; 2005: p. 379) states that there are no hard linguistic criteria for distinguishing codeswitching from borrowing since they are outputs of the same transfer type involving recipient language agentivity. In contrast to the above, it is important to remark that, unlike lexical and grammatical borrowing, codeswitching refers above all to discourse and interaction (Auer 1998). This implies that, in choosing a give language, speakers tend to evaluate the markedness of their potential choices (Myers-Scotton 1993b) and accordingly emphasize instances of codeswitching through a number of linguistic means such as prosody (Manfredi et al. 2014).

When the greater part of a speech community is bilingual, codeswitching may occur extensively. Moroccan Arabic–French codeswitching is a well-studied case in point (Heath 1989). In 'Determiner phrase: how specific is it in Moroccan Arabic-French Codeswitching?' Karima Ziamari investigates nominal insertions in Moroccan Arabic-French codeswitching. As a number of studies have shown, French NPs are embedded in a larger constituent *together with their determiners,* and are further headed by the Arabic determiners *wāḥəd* and *hād*. Using the Matrix Language Frame model (Myers-Scotton 1993a, 2002) and on the basis of an extensive oral corpus, Ziamari seeks to elucidate the motivation behind this unexpected behavior, arguing that morphosyntactic structure alone cannot do justice to explaining the phenomenon, and proposing to take into account the semantic, pragmatic and enunciative mismatch between Moroccan Arabic and French in definiteness, gender and number.

The impact of modern technologies upon centuries-old patterns of contact is explored by Dénes Gazsi in the article 'From Arabia to Persia and back: Arabic-Persian codeswitching among the Al 'Ali tribe in the UAE and Iran'. The article is a thorough analysis of Arabic-Persian codeswitching and the phonological and lexical outcomes of language contact among members of a tribe scattered between the UAE and the coastal Hurmuzgān Province of Iran. Both bilingualism and multidialectalism are at play in the linguistic environment of the speakers, which boasts Modern Standard Arabic, Gulf Colloquial Arabic, Modern Standard Persian, Colloquial Persian and two Persian dialects. The study draws on recorded data with speakers in the UAE and their conversation threads with Iranian tribe members

on social media sites. The results evidence how language choices are determined by the topic of the conversation, the interlocutors' identity and their relationship to each other, all resulting in complex patterns of situational and transactional codeswitching, both inter- and intra-sententially.

The last two contributions tackle the issue of the nexus between language contact and language ideologies. While the linguistic outcomes of language contact involving Arabic have been analyzed from a wide array of perspectives, the study of language attitudes and ideologies lying behind contact-induced phenomena in Arabic and from Arabic have scarcely been studied. The question can therefore be raised whether contact-induced change in and from Arabic unveils different social structures and which effects it has on the social categorization of the language structures in contact.

In asymmetric contact situations such as that of Palestinian Arabic with Israeli Hebrew (Horesh 2015), language ideologies may strongly affect the outputs of language contact. This is illustrated by Nancy Hawker in 'Arabic borrowing of the Hebrew word *menahēl* 'manager': Articulations and ideologies'. The study describes the pragmatic functions – informative or humorous – of the Israeli Hebrew word *menahēl* 'boss' borrowed into Palestinian Arabic. Linguistic anthropology and ethnography can deepen our understanding of contact by introducing language into the 'materiality of ideology' (Grossberg 1986). The case of the Israeli Hebrew borrowing *menahēl* 'boss' in Palestinian Arabic provides material for an analysis that cannot but incorporate the ideologies that represent relations with the conditions of life in the Palestinian-Israeli context. These conditions include economic precariousness for Palestinian day-migrant workers and military control over access and movement in the Occupied Palestinian Territories, coupled by limited autonomy for Palestinian institutions. Palestinians face these conditions with a mixture of stoicism and nationalism, while negotiating with Israeli securitism and economic liberalism. It might seem too obvious to apply this explanatory model to *menahēl*, in its semantic field of power relations, but it is a place to start, not least so as not to annoy 'the boss'.

Valentina Serreli's 'Contact-induced change from speakers' perspectives: a study of language attitudes in Siwa' constitutes a qualitative analysis of the metalinguistic dimension of language contact between Arabic and Siwi, the easternmost Berber language spoken in the Egyptian oasis of Siwa. Siwi has been extensively exposed to contact with different varieties of Arabic (Souag 2014), the most intrusive of which is Egyptian Colloquial Arabic. Nonetheless, geographical and social isolation favored the maintenance of Siwi, which remains the major language of intragroup communication. By combining insights from discourse analysis and contact linguistics, the study seeks to demonstrate that the speakers' perception of

contact-induced change is affected by a high degree of individual variation induced by the differentiated effects of the ongoing economic and social change in Siwa.

6.   Envoy

In the last few decades, the study of language contact emerged from the historical linguistic viewpoint in which it has been traditionally confined and it is now undergoing a process of conceptual renewal and theoretical reconstruction (Nicolai 2007). The present volume joins this new wave of contact linguistics by bringing together leading scholars who address a variety of topics related to contact-induced change, contact languages, codeswitching and language ideologies. It offers, we believe, important insights from different theoretical approaches in connection with other research fields such as descriptive linguistics, sociolinguistics, ethnolinguistics and language acquisition. Over and above all this, the present volume intends to stress the centrality of language contact for Arabic linguistics, and to reveal the significance of Arabic for a multifaceted understanding of language contact.

References

Aikhenvald, A. (2007). Grammars in contact: a cross-linguistic perspective. In A. Aikhenvald and R. M. W. Dixon (eds.), *Grammars in Contact: A CrossLinguistic Typology*. Oxford: Oxford University Press. 1–66.

Arnold, W. (2007). Arabic grammatical borrowing in western Neo-Aramaic. In Y. Matras and J. Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin – New York: Mouton de Gruyter. 185–195.

Auer, P. (1998). From Codeswitching via Language Mixing to Fused Lects: Toward a Dynamic Typology of Bilingual Speech. *The International Journal of Bilingualism* 4 (3): 309–332. https://doi.org/10.1177/13670069990030040101

Bakker, P. and Y. Matras (2013). Introduction. In P. Bakker and Y. Matras (eds.), *Contact Languages. A Comprehensive Guide*. Berlin – New York: Mouton de Gruyter. 1–14. https://doi.org/10.1515/9781614513711.1

Benazzo, S. (2003). The interaction between the development of verb morphology and the acquisition of temporal adverbs of contrast: A longitudinal study in French, English and German L2. In C. Dimroth and M. Starren (eds.), *Information structure and the dynamics of language acquisition*. Amsterdam – Philadelphia: John Benjamins. 187–210. https://doi.org/10.1075/sibil.26.09ben

Chamoreau, C. and I. Léglise (2012). A multi-model approach to contact-induced language change. In C. Chamoreau and I. Léglise (eds.), *Dynamics of Contact-Induced Language Change*. Berlin – New York: Mouton de Gruyter. 1–15. https://doi.org/10.1515/9783110271430.1

Clyne, M. (2003). *Dynamics of Language Contact*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511606526

Coghill, E. (2015). Borrowing of verbal derivational morphology between Semitic languages: the case of Arabic verb derivations in Neo-Aramaic. In F. Gardani, P. Arkadiev and N. Amiridze (eds.), *Borrowed Morphology*. Berlin – New York: Mouton de Gruyter.

Comrie, B. (2011). Creoles and language typology. In C. Lefebvre (ed.), *Creoles, their Substrates and Language Typology*. Amsterdam – Philadelphia: John Benjamins. 599–610. https://doi.org/10.1075/tsl.95.33com

Contini, R. (1999). Le substrat araméen en neo-libanais: Préliminaires à une enquête sisthématique. In M. Lamberti and L. Tonelli (eds.), *Afroasiatica Tergensia: Paper from the 9th Italian Meeting of Afro-Asiatic (Hamito-Semitic) Linguistics*. Padova: Unipress. 101–128.

Diem, W. (1979). Studien zur Frage des Substrats im Arabischen. *Der Islam* 48: 12–80.

Grossberg, L. (1986). On Postmodernism and Articulation: An Interview with Stuart Hall. *Journal of Communication Inquiry* 10(2): 45–60. https://doi.org/10.1177/019685998601000204

Hasplemath, M. (2009). Lexical borrowing: Concepts and issues. In M. Haspelmath and U. Tadmor (eds.). *Loanwords in the World's Languages: A Comparative Handbook*. Berlin – New York: Mouton de Gruyter. 35–54. https://doi.org/10.1515/9783110218442.35

Heath, J. (1989). *From Code-Switching to Borrowing: a case study of Moroccan Arabic*. London – New York: Kegan Paul International.

Heine, B. and T. Kuteva (2003). On contact-induced grammaticalization. In *Studies in Language* 27 (3): 529–572.

Heine, B. and T. Kuteva (2005). *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511614132

Horesh, U. (2015). Structural change in Urban Palestinian Arabic induced by contact with Modern Hebrew. In A. M. Butts, *Semitic Languages in Contact*. Leiden – Boston: E. J. Brill. 198–233. https://doi.org/10.1163/9789004300156_012

King, R. (2000). *The Lexical Basis of Grammatical Borrowing. A Prince Edward Island Case Study*. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.209

Klein, W. and C. Perdue (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13: 301–347. https://doi.org/10.1191/026765897666879396

Kossmann, M. (2013a). Borrowing. In Jonathan Owens (ed.), *The Oxford Handbook of Arabic Linguistics*. Oxford: Oxford University Press. 349–368.

Kossmann, M. (2013b). *The Arabic Influence on Northern Berber*. Leiden – Boston: E. J. Brill. https://doi.org/10.1163/9789004253094

Labov, W. (2007). Transmission and diffusion. *Language* 83: 344–387. https://doi.org/10.1353/lan.2007.0082

Lucas, C. (2012). Contact-induced grammatical change. Towards an explicit account. *Diachronica* 29 (3): 275–300. https://doi.org/10.1075/dia.29.3.01luc

Lucas, C. (2014). Contact-induced language change. In C. Bowern and B. Evans (eds.), *The Routledge Handbook of Historical Linguistics*. London: Routledge. 519–536.

Lucas, C. and E. Lash (2010). Contact as catalyst: The case for Coptic influence in the development of Arabic negation. *Journal of Linguistics* 46 (2): 379–413. https://doi.org/10.1017/S0022226709990235

Manfredi, S. (2013). Native and non-native varieties of Arabic in an emerging urban centre of western Sudan. Evidence from Kadugli. In M. Lafkioui (ed.), *African Arabic: Approaches to Dialectology*. Berlin – New York: Mouton de Gruyter. 13–51. https://doi.org/10.1515/9783110292343.13

Manfredi, S. (2014). Arabic borrowings in Laggorí (Eastern Daju). In T. C. Schadeberg and R. M. Blench (eds.), *Nuba Mountain Language Studies*. Köln: Rüdiger Köppe Verlag. 463–484.

Manfredi, S. (2018). Arabic as a contact language. In R. Bassiouney and A. Benmamoun (eds.), *The Routledge Handbook of Arabic Linguistics*. London: Routledge. 407–420.

Manfredi, S. and M. Tosco (eds.) (2014). *Arabic-based pidgins and creoles*. Special issue of the Journal of Pidgin and Creole Languages 29 (2).

Manfredi, S., M. C. Simeone-Senelle and M. Tosco (2015). Language contact, borrowing and codeswitching. In Amina Mettouchi, Martine Vanhove and Dominique Caubet (eds.), *Corpus-based Studies of Lesser-described Languages: the CorpAfroAs Corpus of Spoken Afro-Asiatic Languages*. Amsterdam: John Benjamins. 283–308.

Matras, Y. (2001). Language Contact. In J. Verschueren, J. O. Östman (eds.), *Handbook of Pragmatics Online*. Amsterdam: John Benjamins.

Matras, Y. (2007). Grammatical borrowing in Domari. In Y. Matras and J. Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin – New York: Mouton de Gruyter. 151–164.

Matras, Y. (2012). *A Grammar of Domari*. Berlin – New York: Mouton de Gruyter. https://doi.org/10.1515/9783110291421

Matras, Y. and P. Bakker (2003). The study of mixed languages. In Y. Matras and P. Bakker (eds.), *The Mixed Languages Debate*. Berlin – New York: Mouton de Gruyter. 1–21.

Matras, Y. and J. Sakel (2007). Introduction. In Y. Matras and J. Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin – New York: Mouton de Gruyter. 1–13.

Miller, C., E. Al-Wer, D. Caubet and J. C. E. Watson (eds.) (2007). *Arabic in the City. Issues in Dialect Contact and Language Variation*. London: Routledge.

Muysken, P. (2000). *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, C. (1993a). *Duelling languages: grammatical structure in codeswitching*. Oxford: Clarendon Press.

Myers-Scotton, C. (1993b). Common and Uncommon Ground: Social and Structural Factors in Codeswitching. *Language in Society* 22 (4): 475–503. https://doi.org/10.1017/S0047404500017449

Myers-Scotton, C. (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299530.001.0001

Nicolai, R. (2007). Language Contact: A Blind Spot in 'Things Linguistic. *Journal of Language Contact* 1: 11–21. https://doi.org/10.1163/000000007792548314

Owens, J. (2001). Creole Arabic: the orphan of all orphans. *Anthropological Linguistics* 43 (3): 348–378.

Ratcliffe, R. R. (2005). Bukhara Arabic: A metatypized dialect of Arabic in Central Asia. In É. Ágnes Csató, B. Isaksson and C. Jahani (eds.), *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*. New York: Routledge. 141–159.

Ross, M. (2006). Metatypy. In K. Brown (ed.), *Encyclopedia of Language and Linguistics*. Amsterdam: Elsevier. 95–99. https://doi.org/10.1016/B0-08-044854-2/00199-1

Ross, M. (2007). Calquing and metatypy. *Journal of Language Contact* 1 (1): 116–143. https://doi.org/10.1163/000000007792548341

Rouchdy, A. (ed.) (2002). *Language Contact and Language Conflict in Arabic*. New York: Routledge.

Savà, G. and M. Tosco (2008). *Ex Uno Plura:* The uneasy road of the Ethiopian languages towards standardization. In *Ausbau* and *Abstand* Languages: Traditional and New Approaches. International Journal of the Sociology of Language 191: 111–139.

Simeone-Senelle, M. C. (1999). L'arabe véhiculaire parlé en Érythrée sur la côte sud de la Mer Rouge de Massawa à Rahayta. *Oriente Moderno* XIX (LXXX): 153–181.

Smits, C. (1998). Two models for the study of language contact: a psycho-linguistic perspective versus a socio-cultural perspective. In M. S. Schmid, J. R. Austin and D. Stein (eds.), *Historical Linguistics 1997*. Selected Papers from the 13th International Conference on Historical Linguistics. Amsterdam: John Benjamins. 377–390. https://doi.org/10.1075/cilt.164.24smi

Souag, L. (2014). *Berber and Arabic in Siwa (Egypt): A Study in Linguistic Contact*. Köln: Rüdiger Köppe.

Taine-Cheikh, C. (2008). Arabe(s) et berbère en contact: le cas mauritanien. In M. Lafkioui (ed.), *Berber in Contact: Linguistic and Sociolinguistic Perspectives*. Köln: Rüdiger Köppe. 113–138.

Thomason, S. G. (2001). *Language contact. An introduction*. Edinburgh: Edinburgh University Press.

Thomason, S. G. (2006). Language contact. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich, A. Zaborski (eds.), *Encyclopaedia of Arabic Language and Linguistics*, vol. 2. Leiden – Boston: E.J. Brill. 664–674. https://doi.org/10.1016/B0-08-044854-2/01901-5

Thomason S. G. and T. Kaufman (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkley: University of California Press.

Tosco, M. (2015). Arabic, and a few good words about empires (but not all of them). In Christel Stolz (ed.), *Language Empires in Comparative Perspective*. Berlin – New York: de Gruyter. 17–39.

Tosco, M. and S. Manfredi (2013). Pidgins and Creoles. In Jonathan Owens (ed.), *The Oxford Handbook of Arabic Linguistics*. London: Oxford University Press (2013) 495–519.

Van Coetsem, F. (1988). *Loan Phonology and the Two Transfer Types in Language Contact*. Dordrecht: Foris.

Van Coetsem, F. (1995). Outlining a Model of the Transmission Phenomenon in Language Contact. In *Leuvense Bijdragen* 84: 63–85.

Vanhove, M. (2012). Roots and patterns in Beja (Cushitic): the issue of language contact with Arabic. In M. Vanhove, T. Stolz and A. Urdze (eds.), *Morphologies in Contact*. Berlin: Akademie Verlag. 311–326. https://doi.org/10.1524/9783050057699.309

Versteegh, K. (2001). Linguistic Contact between Arabic and Other Languages. *Arabica* 48: 470–508.

Versteegh, K. (2010). Contact and the development of Arabic. In Raymond Hickey (ed.), *Handbook of Language Contact*, Chichester: Wiley-Blackwell. 634–651. https://doi.org/10.1002/9781444318159.ch31

Versteegh, K. (2015). An empire of learning: Arabic as a global language. In C. Stolz (ed.), *Language Empires in Comparative Perspective*. Berlin – New York: Mouton de Gruyter. 41–53.

Versteegh, K. (2017). The myth of the mixed languages. In B. Saade and M. Tosco (eds.), *Advances in Maltese Linguistics*. Berlin – New York: Mouton de Gruyter. 217–238.

Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. Berlin – New York: Mouton de Gruyter.

Winford, D. (2003). *An Introduction to Contact Linguistics*. Oxford: Backwell.

Winford, D. (2005). Contact-induced changes. Classification and processes. *Diachronica* 22 (2): 373–427. https://doi.org/10.1075/dia.22.2.05win

Winford, D. (2007). Some issues in the study of language contact. *Journal of Language Contact* 1: 22–39. https://doi.org/10.1163/000000007792548288

Winford, D. (2008). Processes of creole formation and related contact-induced language change. *Journal of Language Contact* 2: 124–148. https://doi.org/10.1163/000000008792525228

Winford, D. (2013). Social factors in contact languages. In P. Bakker and Y. Matras (eds.), *Contact Languages. A Comprehensive Guide*. Berlin – New York: Mouton de Gruyter. 363–416. https://doi.org/10.1515/9781614513711.363

Wohlgemuth, J. (2009). *A Typology of Verbal Borrowings*. Berlin – New York: Mouton de Gruyter. https://doi.org/10.1515/9783110219340

The Arabic component in Domari

Bruno Herin

The goal of this paper is to discuss the Arabic component in the northern dialects of Domari spoken in Lebanon, Syria and Southern Turkey and see to what extent it differs from the Arabic component found in southern Domari, spoken in Jordan and Palestine and already discussed by Matras (2007, 2012).

1.   Background

Domari is an archaic Central Indo-Aryan language spoken by the Dom. It is known to be spoken in Turkey, Syria, Lebanon, Jordan and Palestine. The language is endangered in most places. Syria was the only country where intergenerational transmission seemed in best shape but since the civil war started in 2011, most of the Syrian Dom left to neighbouring countries and more recently to Western Europe.

Apart from various lists collected by western travellers throughout the 19th century, the first full-length description of the variety spoken in Palestine is Macalister (1914). Original fieldwork was carried out by Yaron Matras in the 1990's and 2000's in Jerusalem and resulted in the publication of the first comprehensive grammar of the Jerusalem dialect (Matras 2012). The variety of Aleppo is sketched in Herin (2012), and other published data can be found in Herin (2014: p. 2016).

Data collection by the present author has started in 2009 and is still on-going. It consists of an audio corpus of more than 25 hours of various speech genres ranging from elicitation, dialogues and narratives to oral literature.[1] Places where data were collected are shown in Map 1. Large arrows represent places that I visited and small arrows represent locations for which I collected material from speakers in Beirut, Tripoli and Marj, close to the Syrian border. The area where Domari is spoken is shown roughly in Map 2. The language does not seem to be spoken east of Urfa in

---

[1]   The transcription system used in the present paper follows the rules of the Zeitschrift für Arabische Linguistik, largely based on DIN 31635 and ISO 233-2 except for [χ] represented here with {x}.

southern Turkey anymore. This was first suggested to me by well-informed local observers such Kemal Vural Tarlan, a photographer from Gaziantep who visited most Gypsy communities in Turkey and Mustafa Karabulut, the head of the Dom community in Hatay. Confirmation came when I visited the Dom communities of Diyarbakir and Nusaybin. Although they claimed to speak 'domca' (Dom language in Turkish) or Domani, their speech turned out to be a mixed variety with a Kurdish grammar and a lexicon that partially draws on Domari. In Nusaybin, language attrition and shift to Kurdish were so advanced that only some elderly individuals were able to retrieve a handful of words and sentences in this mixed variety.
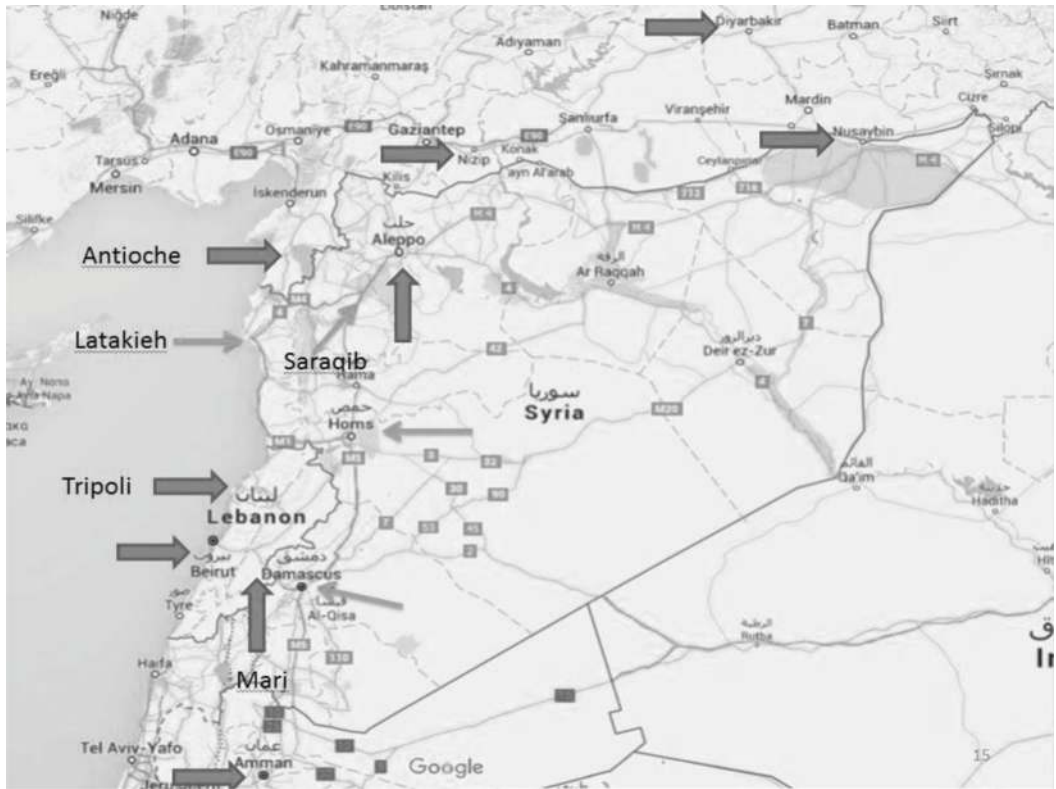
Map 1. Investigated locations (Herin)

All the speakers of Domari living in Arab countries are at least bilingual in Domari and Arabic. The Hatay province (Alexandretta), at least until France decided to detach it from Syria and give it to Turkey in 1939, was for the most part inhabited by ethnic Arabs. Arabic was therefore the most commonly spoken language in the province until very recently and is still actively used amongst Dom above 40 year old in Hatay.

Elsewhere in Turkey, Domari speakers are all fluent in Turkish and also in Kurdish for those living further to the east.
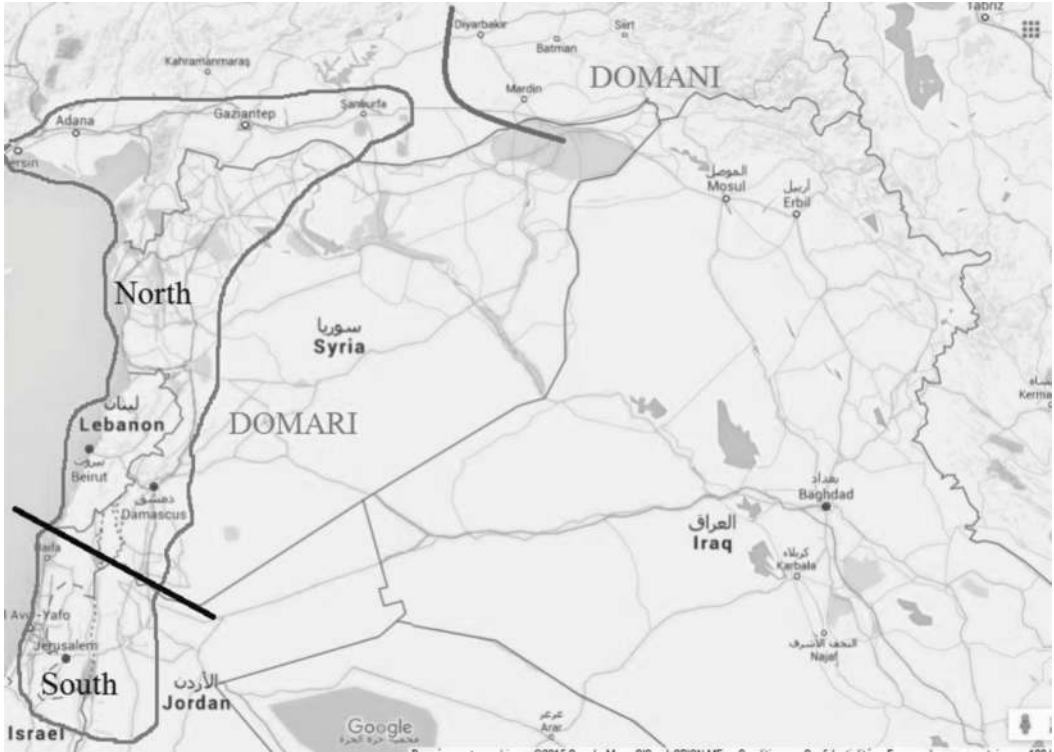
Map 2. Northen-Southern Domari and Domani (Herin)

Domari, besides being almost unknown to most of the scholarly community, is extremely interesting for linguists mostly for two reasons. First, it is surprisingly archaic within the Indo-Aryan family and as such is of great value for historical linguistics. Second, the language has kept important layers of former and current contact languages, starting from Persian, Kurdish, Turkish and eventually Arabic. Domari therefore constitutes a living laboratory for the study of contact phenomena.

Two dialectal areas have so far been identified: southern Domari, spoken in Palestine and Jordan and northern Domari, spoken in Lebanon, Syria and Turkey (see Map 2). Mutual intelligibility is rather low between the two groups. The most salient distinction lies in the maintenance or loss of grammatical gender. Whereas the southern dialects maintained the inherited Indo-Aryan masculine/feminine distinction, the system is severely eroded in the north where the only trace of gender left is a moribund derivational process and a loose agreement pattern (Herin 2016). Another isogloss within northern Domari has to be drawn between the dialects of Beirut and Damascus on one side (henceforth BD) and the dialects spoken in northern Syria and southern Turkey on the other (henceforth ST). The main innovations of the Beirut-Damascus dialect are the passage from [q] to [ʔ] and the loss of the differential subject marker *-ən(ạ)*.

The Arabic component in Palestinian Domari is discussed at length in Matras (2012: p. 368–390). A striking observation is that the distribution of the Arabic material is largely uneven between what Matras documented in the Palestinian variety and what was found in the various northern dialects. In consequence, it appears that the speakers of northern Domari and the speakers of southern Domari must have split long ago and before they ventured into Arabic speaking territories. We can therefore postulate the existence of a pre-split Domari, probably spoken in Anatolia some centuries ago. All these elements prompted me to reassess the Arabic component in Domari in the light of the data collected in all of the locations mentioned above.

Jerusalem Domari is unique in several ways. It is probably the most endangered variety. The number of fluent speakers does not exceed a couple of dozens. The impact of Arabic is so great that it looks fair to state that metatypy, as defined by Ross (1999: p. 7) as a 'change in morphosyntactic type and grammatical organization (and also semantic patterns) which a language undergoes as a result of its speakers' bilingualism in another language', is complete. The most striking features of Jerusalem Domari as far as contact with Arabic is concerned are, amongst others, the wholesale adoption of Arabic connectors (conjunctions, relativiser, fillers …), numbers above five, comparative, core Arabic prepositions, and the auxiliaries *kān*, *ṣār*, *baqa*, *ḍall*, *xallī-* and *bidd-*. The following Example (1) in Jerusalem Domari (Matras 2012: p. 395) is illustrative of the extent of the impact of Arabic:

(1) *ṣāru*   *farružhoñdi* *ʿan*   *ehe*   *raqqāṣīn-an-ta*
     become.PFV.3PL look.IPFV.3PL from DEM.PL dancers-OBL.PL-DAT
     'They started to watch these dancers'

The equivalent in Jerusalem Arabic would probably be (2):

(2) *ṣāru*   *yitfarražu*   *ʿala* *hadōl* *il-raqqāṣīn*
     become.PFV.3PL look.IPFV.3PL on   DEM.PL   DEF-dancers

A quick glance suffices to notice that the non-Arabic elements are the light verb *hōndi*, the demonstrative *ehe* and the case markers *-an-ta*. Of particular interest is the use of the auxiliary *ṣār* inflected as it normally does in Arabic: *ṣāru* 'they became'. In Arabic, the object of the verb *tfarraž* 'watch' requires to be marked with the preposition *ʿala* 'on'. The speaker here replicates oblique marking of the object but does it in a somewhat idiosyncratic and redundant way. First, *ʿala* is replaced by *ʿan*. This is easily explainable because *ʿan* and *ʿala* are interchangeable in some cases. Second, the speaker also marks the object with the dative suffix *-ta*, equivalent to the Arabic prepositions *ʿala* and *ʿan*, instead of the expected ablative *-ki* which normally functions as a prepositional case. The presence of Arabic *ʿan* and the suffix *-ta* is therefore a clear instance of morphological redundancy and a sign that the original

case system is breaking down. Another redundancy is found in plural marking. The speaker employs the Arabic plural form *raqqāṣ-īn* 'dancers' alongside with the Domari oblique plural marking *-an*. The noun *raqqāṣ* is therefore marked twice for plural. None of the phenomena noted above are found in northern Domari, even in those dialects which exhibit the most influence from Arabic.

There seems to be a broad agreement amongst linguists to grant the fatherhood of modern contact linguistics to Weinreich (1953) who argued that the locus of contact was the bilingual's mind. The field has since significantly expanded with the identification of various types of contact that can yield different outputs. The main parameters are intergenerational transmission and the scale of bi/multilingualism within a community. Creoles and mixed languages usually arise when intergenerational transmission is radically renegotiated (Meakins 2013; Tosco & Manfredi 2013). When normal intergenerational transmission is ensured, the main predictors for the linguistic consequence of contact are the scale and historical depth of bi/multilingualism and the sociolinguistic status of the languages involved. Thomason & Kaufman (1988) were amongst the first to provide a rough scale from casual contact to intense contact whose linguistic output materialises in the gradual integration of various word classes and structures. On the whole, there is now an overall consensus amongst contact linguists that semantic transparency and segmentability are the two main parameters involved in the borrowability of linguistic forms (Field 2002).

As noted by Gardani (2015: p. 3), there has been a proliferation of competing terms amongst authors to denote closely related contact phenomena: borrowing, copying, duplication, transfer, replication, diffusion. One convenient terminological solution first introduced in Sakel (2007) and developed at length in Matras (2009) is to consider two kinds of replication: matter replication and pattern replication. Matter replication refers to the transfer of linguistic material proper, whereas pattern replication refers to the transfer of syntactic and semantic layouts.

2.   The Arabic component in northern Domari

2.1   Matter replication

2.1.1   *Open word classes*
In Dom communities where the main contact language is Arabic, there is a general licence to borrow anything from open word classes (verbs, nouns and adjectives), even when a pre-Arabic option is available. These are items that were in use prior to contact with Arabic and maybe inherited or of Turkish, Kurdish or Persian origin. The degree to which this licence is instantiated in actual speech varies from

speaker to speaker. Speakers for which Arabic is dominant will tend to integrate more Arabic derived lexemes. This is not per se an indicator of lexical knowledge because many speakers are aware of the existence of pre-Arabic options. Some examples are *ruzz* vs *brīnǧ* 'rice', *nəǧma* vs *yēldəz* 'star', *wāṭi* vs *alčāx* 'low', *ʿāli* vs *ūǧa* 'high'. Some other items fell into oblivion and only a handful of speakers are able to retrieve them, such as *sīw* 'apple' or *lōrga* 'tomato' replaced by Arabic derived *təffāḥa* and *bandōra*. Items for which there is no pre-Arabic option, the selection of the Arabic item is of course the default choice. Domari has fully integrated all the segmental phonology of Arabic. All the back consonants peculiar to Arabic (/ġ/, /x/, /ḥ/, /ʿ/ and /h/) and the pharyngealised consonants (/ṭ/, /ṣ/, /ḍ/ and /ẓ/) are also part of the inventory of Domari. The uvular /q/ is found only in the dialects spoken in northern Syria and southern Turkey. In the dialect of Beirut/Damascus, it became /ʾ/. This glottal realisation is of course contact-induced as it is found in most sedentary varieties of Arabic spoken in the northern Levant. Examples are *ʾrī kar* 'read' (< Arabic *q-r-ʾ*), *ʾər* 'son' (< *qər*, derived from Kurdish), *ʾāršōs-ta* 'in front of him' (< Turkish *karşı* 'face'), *ʿaʾʾōr* 'nut' (< *ʿaqqōr*, inherited). The pharyngeals /ḥ/ and /ʿ/ may have been introduced into Domari prior to contact with Arabic because it already appears in loans from Kurmandji Kurdish such as *ʿārḍ* 'earth', *mʿōri* 'ant', *ḥaft* 'seven', *ḥašt* 'eight', although the presence of pharyngeals in some dialects of Kurdish may also be due to contact with Arabic (Haig & Öpengin, forthcoming).

Arabic nouns and adjectives are all integrated into Domari according to their original segmental phonology. The only thing that is modified is stress assignment. In Domari, stress appears on the last syllable: *drōná* 'big', *m(ə)nás* 'man'. Loans from Arabic usually follow this stress assignment pattern: *aǧnabí* 'foreign' (Arabic *áǧnabi*). An interesting phenomenon is that Arabic epenthetic vowels are reinterpreted as plain vowels, and bear stress in final syllables: *waḍáʿ* 'situation' (Arabic *wáḍəʿ*), *ṣaʿáb* (Arabic *ṣáʿ(ə)b*). The feminine morpheme *-a* in Levantine Arabic often undergoes raising to [e] in non-emphatic and non-guttural contexts. These are always integrated into Domari with final [a]: *mádrase* 'school' > *madrasá*.

As noted above, gender is no longer an inflectional category in northern Domari. When Arabic adjectives are borrowed into Domari, they only appear in their masculine forms even whith female referents, as exemplified in (3), where the speaker talks about a girl:

(3) *ādami   e   u   maḥšūm e*
     humane COP and decent   COP
     'She is humane and decent'

Comparative constructions are not uniform across northern dialects. In the dialects of northern Syria and southern Turkey, comparatives are formed using the Kurdish derived suffix *-tar* and superlatives are formed with the Turkish derived

morpheme *ān* followed by the adjective. In the dialect of Beirut/Damascus, these options are not available and speakers simply borrow comparative forms from Arabic. Example (4) illustrates the use of Arabic *azġar* 'smaller', from *zġīr* 'small' as a suppletive comparative form for Domari *tnōta ~ trōta* 'small'. Example (5) illustrates a superlative construction with Arabic *akbar* 'bigger' from *kbīr* 'big' as a suppletion for Domari *drōn(g)a* 'big'.

(4) *amīn lāzim l-pāran   azġar   wēšōma*
     1PL   must   SBJV-we.take smaller 1PL.ABL
     'We have to marry someone younger than us'

(5) *lāfty-an-ma   mā akbar yōkāk ǝštōm*
     girl-OBL.PL-IN 1SG older one   COP.1SG
     'I'm the oldest amongst the girls'

In cases when adjectives for which there are no pre-Arabic equivalent and no available morphological derivation are integrated, Domari employs the same syntactic device as in Arabic with a postponed *aktar* 'more' after the adjective. In (6), the Arabic adjective *mdallal* 'spoiled' does not have an aCCaC derivation so the only way is to use *aktar*. The Domari copula is placed between the adjective and *aktar*.

(6) *mā mdallal ǝštōm   aktar sā-ēn-ki   bābōm-ka*
     1SG spoiled COP.1SG more all-OBL.PL-ABL my.father-AD
     'My father spoils me the most (literally I'm more spoiled than everyone at my father).'

This strategy of borrowing wholesale comparative forms from Arabic was already described by Matras (2012) in Palestinian Domari, for which he coined the expression 'bilingual suppletion'. Bilingual suppletion also extends in Palestinian Domari to the use of Arabic numerals above four when modifying a noun. Matras (2012: p. 192) reports for example *arbaʕ qar-e* 'four donkeys' with Arabic *arbaʕ* 'four' and inherited *qar* 'donkey' marked for plural with *-e*. Most often though, with numerals above three, speakers of Jerusalem Domari favour the use of Arabic numerals and Arabic nouns: *taran dīs* 'three days' but *sabaʕt iyyām* 'seven days' (Matras 2012: p. 194). This kind of bilingual suppletion also occurs in Beirut Domari but only in the speech of speakers for which Arabic is dominant. The speaker who uttered (7) has, according to her own judgment, difficulties to retrieve Domari numerals above six and therefore uses Arabic items. She selects the Arabic numerals *sabǝʕ* 'seven' and the plural noun *iyyām* instead of the pre-Arabic equivalent *ḥaft dīs*.

(7) *māndēnd ya'ni ši   sabǝʕ tiyyām ǧištand āwištande*
     they.stayed that.is some seven days   they.go they.come
     'They kept coming and going for let's say seven days'

It seems therefore that bilingual suppletion, at least in the case of numerals, is closely related to language dominance and linguistic competence. This instable pattern can stabilise if transmitted as such to the next generation. This is what appears to have happened in Jerusalem Domari. One question remains unanswered of course: why use an Arabic nouns instead of the inherited option? Or to put it differently: what is it that makes speakers reluctant to use *sabə' dīs* or *sabə' dīs-ī(n)*? One possibility is that an expression such as *sabə' tiyyām* enjoys a greater level of lexicalisation so there is no rule transfer involved in any kind and the speaker simply imports from Arabic the phrase as it is. One could also claim that since Arabic requires a plural and Domari a singular from 3 to 10, speakers prefer to apply the rules of the dominant language and select an Arabic plural (Souag & Kherbache 2016). The problem is that Arabic nouns are selected even with higher numerals above 10, as shown in (8), from the same speaker. Here she uses *wāḥad u 'išrīn* '21' followed by Arabic singular *sane*, but in the other part of the utterance she uses inherited *trən wars* 'three years'. Here there are no conflicting rules because both Arabic and Domari select the singular form of the noun.

(8) *'əmrōs   wāḥad u   'išrīn   sane akbar 'rōm-ki   b-trən   wars*
      her.age   one   and twenty   year   older my.son-ABL with-three years
      'She is 21 years old, three years older than my son'

It seems therefore that the use of an Arabic numeral implies de facto the use of the Arabic form for the counted item. Unfortunately, the corpus does not contain enough tokens of these bilingual suppletions as it was found only in the speech of two informants (two sisters) for which Arabic is clearly dominant. All the other informants limit their use of Arabic numerals to the expression of time and date. Because these cases of bilingual suppletion are only found in the speech of speakers whose Arabic is dominant, it is quite obvious that they favour this strategy because it allows them to fulfil efficiently the communicative task they are engaged in. For these speakers, trying to retrieve Domari numerals appears to be too costly from a cognitive perspective.

Arabic plural forms are rarely borrowed as such in northern Domari. Speakers usually select the singular and add the Domari plural marker *-ī(n)*: *šēx-īn* 'elders', *ṣaxr-īn* 'rocks', *akl-īn* 'foods', *dars-īn* 'lessons'. Borrowing of Arabic plural forms does occur but only with more lexicalised plurals such as *qarāyib* 'relatives', *ẓrūf* 'conditions', *žīrān* 'neighbours' and also plurals used adverbially such as *awqāt* 'times' or *ayyām* 'days', both used in the sense of 'sometimes'. Example (9) illustrates the use of *qarāyib*, technically plural of *qarīb* but practically an autonomous entry in the Arabic lexicon.

(9) *pārane ʾarāyib-an-ki   bass inno ġarīb-a-ki   nə-dēne*
     we.take relatives-OBL.PL-ABL but   COMP stranger-OBL-ABL NEG-we.give
     'We take from our relatives but we don't give (women) to a stranger'

This situation sharply contrasts with what is observed in Palestinian Domari where Arabic plural forms are always recruited. Matras (2012: p. 340) reports for example the use of the Arabic plural *madāris* 'schools' (10) whereas speakers of northern dialects always select the singular and pluralise it with a Domari suffix, whether *-ī(n)*(11) or oblique plural *-an* if inflected for case.

(10) *na nērdedim   madāris-an-ka*
      NEG they.sent.me schools-OBL.PL-DAT
      'They didn't send me to school'

(11) *madras-īn ašti bass bəməll   əštande*
      school-PL EXS but   expensive COP.3PL
      'There are schools but they are expensive'

As far as the integration of Arabic verbs is concerned, all the dialects of Domari make use of the same light verb strategy. Domari has at its disposal two light verbs: *kar-* 'do' and *h(ō)-* 'become'. Roughly speaking, transitive verbs are integrated using *kar-* and intransitive verbs using *h(ō)-*. Speakers isolate the Arabic imperfective stem and impose a /i/ vocalism on the final syllable: *nsī kar-* 'forget' (Arabic stem *-nsa*), *stannī kar* 'wait' (Arabic stem *-stanna*). Table 1 summarizes how Arabic verbs are integrated according to their derivational template.

Table 1. Integration of Arabic verbs

Measure	Arabic	Domari
I (transitive)	*šakar-yuškur* 'thank'	*škər kar-*
I (intransitive)	*fall-yfill* 'escape'	*fəll h(ō)-*
II	*rabba-yrabbi* 'raise'	*rabbī kar-*
III	*sāʿad-ysāʿid* 'help'	*sāʿəd kar-*
IV	–	
V	*trabba-yitrabba* 'be raised'	*rabbī h(ō)-*
VI	–	
VII	*nǧabar-yinǧibir* 'be forced'	*ǧbər h(ō)-*
VIII	*xtalaf-yixtalif* 'be different'	*xtaləf (h)ō-*
IX	–	
X	*staġrab-yistaġrib* 'to be suprised'	*staġrəb h(ō)-*

No instances of measures IV, VI and IX could be found in the various corpora. This is not surprising for measure IV because it is not productive in most Arabic

dialects speakers of Domari are in contact with. Measure IX is rendered in Domari using the corresponding adjective followed by *h(ō)-* 'become': *lōra (h)rēnd* 'they turned red' (Arabic *ḥmarru*). The absence of measure VI is probably due to its limited productivity in Arabic itself and because inherited options are available in Domari such as *lagīš kar-* 'quarrel' (Arabic *txānaq, ṭṭāwaš, thāwaš, tkāwan, tḥārab*). Interestingly also, speakers of Domari rely on the imperfective stems of measures II and I for measures V and VII. So *rabbī h(ō)-* 'be raised' and not a hypothetical **trabbī h(ō)-*. The same goes for *ğbər h(ō)-* and not **nğabər h(ō)-*. This means that speakers are aware of the derivational link between measure I *ğabar* 'impose' and measure IX *nğabar* 'be imposed' on the one hand and *rabba* 'raise' and *trabba* 'be raised' on the other. Such integration strategy indicates that speakers have an intimate knowledge of Arabic because they are able to extract the imperfective stem and apply a specific vocalic pattern, as well as being aware of the different semantic links between the derivational templates.

2.1.2   *Closed word classes*

As far as closed word classes are concerned, the integration of Arabic prepositions differs substantially between the Beirut/Damascus dialect and the varieties spoken in northern Syria and southern Turkey. In the dialect of Beirut/Damascus, the core Arabic prepositions *b-* 'in, with' and *min* 'from' do occur in specific contexts. In (12), the speaker uses the Arabic sequence *aktar min* 'more than' which seems to have been borrowed as such. In (13), the Arabic structure *qabəl* + Y + *b-* is replicated with the preposition *b-*. In the dialects of northern Syria and southern Turkey, this preposition may surface but only when it refers to an instrumental, in which case it alternates with the inessive case marker *-ma*: *b-črī ~ čəry-a-ma* 'with a knife'. Unlike Palestinian Domari, other core Arabic prepositions such as *ʿind* 'at' and *maʿ* 'with' never occur, neither in BD nor ST.

(12)   *aktər min   wīst   warəs e*
       more from twenty year   COP
       'It's been more than twenty years'

(13)   *ʾabl   əb-dīs-ā   āndōs-sa   wāšī*
       before in-day-INDF he.brought-3PL 3SG.COM
       'He brought them with him one day before'

As noted above, Palestinian Domari borrowed many Arabic auxiliaries (*kān, ṣār, baqa, ḍall, xallī-* and *bidd-*) and inflects them as in the source language, as shown in (1). Amongst these, only *kān* and *ṣār* occur the Beirut/Damascus dialect, but with an important difference: they do not inflect. In (14a), the speaker translates Arabic *ṣāru yšūfu ḥāl-hum* 'they started showing off' (lit. 'see themselves') but *ṣār* remains uninflected. Dialects of northern Syria and southern Turkey did not replicate any

Arabic auxiliaries. What they did instead is to resort to pattern replication in the case of inchoative *ṣār* (14b). Since its primary meaning in Arabic is 'become', they selected the corresponding Domari verb *hr-* and replicated the Arabic structure *ṣār* + subjunctive. In (15), the use of *kān* locates the event in the past and gives it an iterative/habitual aspect. The imperfective of *kān* was also replicated, as shown in (16). It has the same semantics as in Arabic, describing a possible state of affairs not attested at the time of utterance, but like *ṣār* and *kān* it remains invariable. The floating syntax especially of *kān* and *bikūn* suggests that they are not integrated as auxiliaries but rather as predicate modifying adverbs.

(14)   a.   *ṣār*   *l-adkand*   *pā-ēn*
            he.became   SBJV-they.see   REFL-OBL.PL
       b.   *hrēnd*   *l-dakənd*   *pā-ən*
            they.became SBJV-they.see   REFL-OBL.PL
            'They started showing off'

(15)   *ʾawwalma ğīrsāwīrōm nawa ānane   ğāne kān   wēsrēna   xaldē-ma*
       when   I.married   new we.bring we.go he.was we.stayed Khalde-IN
       'After I got married we used to go and stay for a while in Khaldé'

(16)   *bikūn brīnǧ-īn   naʾʾəʿ(h)rēnde*
       he.is   rice-PL   they.are.soaked
       'The rice will be soaked'

Domari does not have any inherited connector. Drawing the exhaustive inventory of the connectors found in Domari is beyond the scope of this paper. I will just mention those which exhibit differences between the Beirut/Damascus dialect on one part and the dialects of northern Syria and southern Turkey on the other. As hinted above, the Arabic relativizer *illi* is commonly used in both groups. Like Arabic, Domari also makes use of the resumptive pronoun strategy. In Example (17) we have a headless relative clause and a pronominal object indexed on the verb. There are indications that *illi* in ST dialects is a recent borrowing because it competes with the Turkish/Kurdish derived relativizer *ki*, still commonly found in almost lexicalised phrases such as *wars ki mḍī(h)ra* 'last year' (year REL it.passed) and also sporadically in proper relative clauses. In this case too, the resumptive pronoun strategy is used. Finally, it should be added that this strategy in Domari need not be necessarily replicated from Arabic as it is also attested in previous contact languages such as Kurdish and Persian.

(17)   *tōnde   lōn ṭaṭōsar alli mangand-s-e*
       they.put salt chilli   REL they.want-OBJ.3SG-PRS
       'They add salt, chilli, whatever they want'

As far conjunctions are concerned BD and ST differ in the extent to which they resort to Arabic conditionals. In BD, only Arabic *iza* (18) and *law* (19) surface whereas in ST, Arabic conditionals compete with the Turkish derived clitic = *sa* or zero-marking.

(18)   *bīnāre   iza māns-as   ǝgzǝrda baḥr-a-ma*
       it.scares if   individual-ACC it.bite   sea-OBL-IN
       'It's scary if it bites someone in the sea'

(19)   *law pārdōm ēr-as   illi   mra   kakī hre-ya   manğōm kān*
       if   I.took   DEM-ACC REL he.died what it.became-PST 1SG.IN   PST
       'If I had married the one who died, what would have happened to me?'

Coordination is another function for which the reliance on Arabic derived materials differs between BD and ST dialects. The original system, interesting from a typological point of view and found in ST, uses different constructions for NP coordination and clausal coordination. NP's are coordinated using the conjunction *la* (*mā la tō* 'me and you') and clauses are coordinated with the enclitic *ši*. Both morphemes are derived from Kurdish. In BD, this system is showing signs of breakdown because it is being replaced by the Arabic conjunctions *w* 'and' for both NP and clausal co-ordinations. Compare for that matter the way Arabic *ḥāl* clauses are replicated in both BD (20a) and ST (20b). They both translate the Arabic sentence *w iḥna ʿam mnistanna ṣārat tiṯliğ* 'As we were waiting, it started to snow', obtained by elicitation.

(20)   a.   *ṣār   l-wār   xīw w   amīn stannī=kištan*
            ṢĀR SBJV-it.hits ice and 1PL   wait=we.do.PROG
       b.   *xīw āyra   amīn ši   akī=kaštinne*
            ice it.came 1PL   and eye=we.do.PROG
            'It started snowing as we were waiting'

It appears clearly that while BD replicates Arabic matter, ST only replicates pattern from Arabic and relies on internal material to copy the Arabic structure *w* + pronoun.

2.2   Pattern replication

2.2.1   *Syntax of the NP*
As notes above, the tendency to rely on matter replication is much more pervasive in the dialect of Beirut/Damascus than in those spoken in northern Syria and southern Turkey, as exemplified by the replication of the Arabic inchoative construction using *ṣār* and *ḥāl* clauses. In the realm of syntax, the impact of Arabic is also uneven between the two groups. As far as the syntax of noun phrases is

concerned, both groups seem to have preserved to a large extent the traditional order modifier + noun. This sharply contrasts with Palestinian Domari which appears to have almost completely adopted the Arabic syntax noun + modifier (Matras 2007: p. 159). In both BD and ST, the most common order is adjective + noun: *bxēz nārṇ* 'good man', *aṣlī gɚr* 'traditional grease', *drōna bārōm* 'my older brother'. Instances of noun + adjective do occur but all the examples were produced during elicitation sessions involving translation from Arabic to Domari: *bārōm ɚdrōna* 'my older brother', *lāfty-ā tnōtī* 'a small girl'. In spontaneous speech, the most common order is overwhelmingly adjective + noun. In genitive constructions, the traditional order is also modifier + modified: *māmōm 'ɚr* 'the son of my uncle, my cousin'. Things can be different in complex genitive construction involving two modifiers, as shown in (21). We see here the second modifier being placed to the right, and not to the left as in the case of the first modifier.

(21) *mām-ō-s   'ɚr   mɚns-ō-m-ki*
      uncle-SG-3SG son husband-SG-1SG-ABL
      'the son of the uncle of my husband (my husband's cousin)'

Matras (2012: p. 169) reports that the constituent order in genitive constructions in Palestinian Domari is almost always modified + modifier as in Arabic, leaving the order modifier + modified rather marginal. Matras (2012: p. 170) also mentions a genitive particle *kāk-* to which bound pronouns suffixes: *pl-e-m kakī-m* 'my money' (money-PL-1SG POSS-1SG). This particle, seemingly cognate with the northern Domari interrogative *kakī* 'what, which', replicates the pan-Levantine genitive particle *tabaʿ*. Such a construction is unattested in both BD and ST.

2.2.2 *Syntax of simplex clauses*

Constituent order in simplex clauses in Levantine Arabic is, depending on information structure, either VSO or SVO. In Beirut Domari, the most common order appears to be SVO, illustrated in (22a). Only one instance of SOV surfaces in the corpus (22b).

(22) a. *bābōm-ki   ʿašīrōs   nɚ-ḥāskand   nāčīs-as*
          my.father-ABL his.clan NEG-they.like dancing-ACC
          'My father's clan doesn't like dancing'
      b. *pānğā dōm   gāl   na-krand*
          3PL   Domari word NEG-they.do
          'They don't speak Domari'

A striking feature of ST is that the order SOV is much more prevalent than in the dialect spoken to the south (BD and Palestinian), as exemplified in (23), recorded

in Aleppo. The SOV order is, as far as one can judge from available data, inexistent in Jerusalem Domari.

(23) *pāvəs   vēštyən   pāṣōm   qaḥwa pyən   vēsrēn   qaḥwa pīrēn*
       come.PL we.sit   1SG.AD   coffee we.drink   we.sat   coffee we.drank
       '(I told them) come and let's sit at my place to drink coffee. We sat down and drank coffee'

The SOV order is also likely to be contact induced or at least, if it ever was the original Domari order, to have been maintained because it is commonly found in neighbouring languages such as Turkish, Kurdish and Persian. A quick look at the map representing the order of subject, object and verb in the World Atlas of Language Structures (Chapter 81A, Dryer 2013) clearly suggests that the SOV order is an areal feature found in Anatolia, the Caucasus, Western Asia and India, irrespective of genetic affiliation. Considering constituent order in both the NP, whether noun-adjective or genitive constructions, and in simplex clauses, it appears that the impact of Arabic materialises in a gradual drift from a head-final to a head-initial syntactical typology on a north-south axis.

2.2.3   *Negation*

One last feature for which the impact of Arabic differs from one group to the other is negation. In the Palestinian variety, Matras (2012: pp. 347–351) besides the inherited negator *na*, reports the use of Arabic *ma* and *mišš* and *la …wala*. Jerusalem Domari also has the pattern *(n-) …-ʾ*, with initial *n-* being optional, similar to Palestinian Arabic *(ma-) …-š*: *(n)-kafikarse-ʾ* 'it's not enough', equivalent to Palestinian Arabic *(ma)-bikaffī-š*. Northern dialects do not exhibit this final glottal element. What they do instead is stress the last syllable of the verb when a imperfective stem is used: *ǧāname* 'I know' vs *n-ǧān(a)mé* 'I don't know'. Both *n-* and final stress are compulsory. The compound negation found in Palestine was therefore not replicated from Arabic since an incipient pattern was part of pre-split Domari. The only thing that is modelled on Palestinian Arabic is the optionality of the first element of the negation. As far as *ma* is concerned, it is not straightforward at all that it is a borrowing from Arabic because it also appears in northern dialects as an inherited Indic morpheme used in the negation of jussive mood. The inherited mood-based complementary distribution of *na-* and *ma-* may have been lost in Palestinian Domari probably due to contact with Arabic and because of the homophony between the two markers. Arabic *miš* is mostly used in non-verbal predication. It appears to have make its way into all southern dialects, as shown in (24a), recorded in Jordan. This is impossible in northern dialects, which only rely on inherited *nnye* 'it is not' placed after the predicate (24b).

(24) a. *hā   zāra   muš gūzal-i*
          DEM boy   NEG good-COP
          'This boy is not good'
      b. *hā   ammat bxēz ənnye*
          DEM people good COP.NEG
          'These people are not good'

The corresponding form of *miš* in most Syrian dialects of Arabic is *mū*. While no instances of *mū* were recorded in any ST dialects, it surfaces in both elicited material and spontaneous speech in BD. Its distribution however is different from what is found in the varieties spoken in Palestine and Jordan. It appears that the use of *mū* is licenced in only two cases. The most common one is elliptic constructions where the negator has scope over one constituent only and not the whole predication as in southern dialects (26a). As far as verbal negation is concerned, *mū* can also appear with verbs in the subjunctive (26b).

(25) *ānande   bāfər bass mū sạ, īsa ē   nawa ğīl*
      they.bring many but NEG all now DEM new   generation
      n-āništar   bāfər
      NEG-it.brings   many
      'They have many children but not all of them, this new generation doesn't have many children'

(26) a. *yōka mū   wēsnār-am*
          one   NEG he.wakes.up.SBJV-1SG
          'Nobody wakes me up!'
      b. *biğūz masalan   mū   māntyar   wāš   məṣrī*
          maybe for.instance NEG it.remains.SBJV 3SG.COM money
          'For example, she might not have any money left'

The Arabic contrastive negative coordination markers *lā …walā* 'neither …nor' are found in all the dialects so far investigated, whether ST, BD or Palestinian (27, recorded in Beirut).

(27) *kānye   lā   'wāṣ   wala lagīšīn*
      EXS.NEG neither shooting nor   fights
      'There are neither shootings nor fights'

However, a construction drawing on inherited material *n- …n-* is still available in ST: *n-amā nə-bēnōm* 'neither me nor my sister'. A mixed construction was also recorded in the speech of one informant from Sarāqib (north-western Syria, see Map 1): *nə-mā wala bēnōm* 'neither me nor my sister'. It is not entirely clear whether

the construction that draws on inherited material is the outcome of pattern replication from Arabic or was available prior to contact with Arabic.

3.   Conclusion

The primary goal of this paper is descriptive because Domari remains poorly documented. The need to bring to light first hand unpublished linguistic data about a language the scholarly community knows so little about is therefore urgent, especially in the light of the recent dramatic developments Syria has witnessed, where virtually all Dom communities are now displaced, putting even more pressure on language transmission. As noted above, Domari is of particular interest to linguists studying contact phenomena because its speakers have been since time immemorial at least bilingual. Multilingualism has at all times concerned all the speakers of the community and for long stretches of time for each contact language. Beside the core central Indic component, various 'foreign' layers are identifiable starting from Dardic in north western India, Persian, Kurdish, Turkish and eventually Arabic. This latter layer is particularly worth investigating because it is not uniform across the dialectal groups that have been identified: southern Domari spoken in Palestine and Jordan, and northern Domari spoken in Syria, Lebanon and southern Turkey which itself subdivides into the Beirut/Damascus dialect and the varieties of northern Syria and southern Turkey. It was shown above that the Arabic component of Palestinian Arabic was far from being shared with other dialects. Bilingual suppletion found in Palestine for comparatives and numerals above four is unknown in ST and present in BD with comparatives and incipient with numerals only amongst the most Arabicized speakers; Arabic plurals common in Palestine are extremely restricted in northern Domari; core Arabic prepositions found in Jerusalem Domari rarely appear in the north; Arabic auxiliaries are either replicated according to pattern or non-inflected matter, as opposed to Palestinian Domari which relies entirely on matter replication and also borrows Arabic inflections; relativisation and conditionals are rather uniform cross-dialectally but recently replaced Kurdish/Turkish morphemes in the north; the syntax of NP's is still largely head-final in the north but underwent convergence with Arabic in the south; convergence in the syntax of simplex clauses in complete in the south, almost so in BD whereas ST dialects still exhibit to a fair extent the SOV order and finally, while Palestinian Domari closely resembles Arabic in terms of both matter and pattern, matter replication is highly restricted in the north. The picture that arises is therefore rather straightforward. Convergence towards Arabic is gradual from north to south, with at one end Palestinian Arabic which relies heavily on matter replication and at the other end the dialects spoken in northern Syria and

southern Turkey in which the integration of Arabic matter is much more restricted and reliance of pattern replication favoured in many cases. This also suggests that phenomena such as bilingual suppletion in particular and large scale transfer of matter in general involve a greater historical depth of bilingualism and a more advanced stage of language attrition.

List of abbreviations

ABL	ablative	EXS	existential	POSS	possessive
ACC	accusative	IN	inessive	PROG	progressive
AD	adessive	INDF	indefinite	PRS	present
COM	comitative	IPFV	imperfective	PST	past
COMP	complementizer	NEG	negation	REFL	reflexive
COP	copula	OBJ	object	REL	relative
DAT	dative	OBL	oblique	SBJV	subjunctive
DEF	definite	PFV	perfective		
DEM	demonstrative	PL	plural		

References

Dryer, M. (2013). Order of Subject, Object and Verb. In M. Dryer & M. Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/81, Accessed on 2016-06-09.)

Field, F. (2002). *Linguistic Borrowing in Bilingual Contexts*. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.62

Gardani, F., P. Arkadiev and N. Amiridze (2015). Borrowed morphology: An overview. In Gardani, F., P. Arkadiev and N. Amiridze (eds.), *Borrowed Morphology*. Berlin, Boston: Mouton de Gruyter. 1–23.

Haig, G. and E. Öpengin.(forthcoming). Kurmanji Kurdish in Turkey: Structure, varieties, and status. In C. Bulut (ed.), *Linguistic Minorities in Turkey and Turkic Speaking Minorities of the Peripheries*. Wiesbaden: Harrassowitz.

Herin, B. (2012). The Domari language of Aleppo (Syria). *Linguistic discoveries* 10(2): 1–52.

Herin, B. (2014). The northern dialects of Domari. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 164: 407–450.

Herin, Bruno. (2016). Elements of Domari dialectology. *Mediterranean Language Review* 23: 33–73.

Macalister, R. A. S. (1914). *The Language of the Nawar or Zutt, the Nomad Smiths of Palestine*. London: Edinburgh University Press.

Manfredi, S., M.-C. Siméone-Senelle, and M. Tosco (2015). Language contact, borrowing and codeswitching. In A. Mettouchi, M. Vanhove and D. Caubet (eds.), *Corpus-based Studies of lesser-described Languages: the CorpAfroAs Corpus of spoken AfroAsiatic languages*. Amsterdam: John Benjamins. 283–308.

Matras, Y. (2007). Grammatical borrowing in Domari. In Y. Matras and J. Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter. 151–164.

Matras, Y. (2009). *Language Contact*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511809873

Matras, Y. (2012). *A Grammar of Domari*. (Mouton Grammar Library). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110291421

Meakins, F. (2013). Mixed languages. In Y. Matras and P. Bakker, *Contact Languages: A Comprehensive Guide*. Berlin: Mouton de Gruyter. 159–228. https://doi.org/10.1515/9781614513711.159

Ross, M. (1999). *Exploring metatypy: how does contact-induced typological change come about? Paper presented to the Annual Meeting of the Australian Linguistics Society* . Perth.

Sakel, J. (2007). Types of loan: Matter and pattern. In Y. Matras and J. Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective* Berlin: Mouton de Gruyter. 15–29.

Souag, L. and F. Kherbache (2016). Syntactically conditioned code-switching? The syntax of numerals in Beni-Snous Berber. *International Journal of Bilingualism* 20(2): 97–115.

Thomason, S. and T. Kaufman. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Tosco, M. and S. Manfredi. (2013). Pidgin and creoles. In J. Owens (ed.), *The Oxford Handbook of Arabic Linguistics*. Oxford: Oxford University Press. 495–519.

Weinreich, U. (1953). *Languages in Contact: Findings and Problems*. The Hague: Mouton de Gruyter.

Syntactic outcomes of contact in Sason Arabic

Faruk Akkuş and Elabbas Benmamoun

University of Pennsylvania / Duke University

In this paper, we discuss a number of morphosyntactic properties of Sason Arabic, which could be strongly argued to be due to contact with the neighboring languages, some of which have head-final properties. We argue that Sason Arabic patterns with both its Arabic neighbors and the typologically different surrounding, and sociolinguistically dominant languages, particularly Kurdish and Turkish. We aim to show that syntactic constructions in contact contexts can provide important insights into the nature of the contact and the history of the language and its speakers.

Keywords: Sason Arabic, language contact, copula, (in)definiteness marking

1. Introduction

Contact between languages is as old as human history. All human activities of different scales, such as conflict, population movement, and trade, among others, result in contact between languages and competition for linguistic space with all attendant consequences for all languages involved, though to varying degrees of intensity and impact. Sustained contact can lead to changes in the linguistic system, its sound inventory and patterns, its lexicon, word structure, and syntax. In the Semitic language context, it is not surprising that Akkadian and Amharic display head final properties and thus differ from most of their Semitic counterparts, such as Arabic, particularly the varieties spoken in the Arab world, and Hebrew. Akkadian was in contact with Sumerian, a head final non-Semitic language, and Amharic is in contact with Cushitic languages which are head final. It would be appropriate to hypothesize that the head final properties found in Akkadian and Amharic could mostly likely be due to contact, rather than internal change within the two languages triggered by some language internal pressures. This paper focuses on a dialect of Arabic that has been isolated from the Arab world for centuries and that has not been used as a language of literacy. Sason Arabic, spoken in Turkey, is located in an area where it is not the dominant language.

We discuss several contact-induced changes in Sason Arabic, henceforth SA, which patterns with both its Arabic neighbors, particularly the so-called Mesopotamian varieties (such as the Iraqi variety/varieties of Mosul) and the neighboring languages that are typologically different, particularly Kurdish and Turkish. We advance the thesis that language contact with the typologically different neighboring languages has led to significant morphosyntactic changes besides the lexical influences discussed in Talay (2001), Isaksson (2005), Jastrow (2006a), and Lahdo (2009).

The morphosyntactic properties we will focus on are indefiniteness, light verb constructions, periphrastic causatives, negation and copula constructions. All these constructions reveal significant syntactic changes that we can confidently attribute to contact.

The paper is organized as follows: In Section 2, we provide a brief description of Sason Arabic and some of its morphological properties that have emerged as a result of language contact. The section also introduces the general syntactic features of the language, which pattern like other Arabic varieties. Section 3 discusses the contact-induced morphosyntactic changes and gives an account of the clause structure of SA.

2.   Sason Arabic

Sason Arabic is one of several Arabic varieties spoken in Anatolia and which are part of the larger Mesopotamian dialect area. They are typically considered to be close to Iraqi Arabic dialects. Jastrow (1978, 2006a) groups the Sason dialect with the other members of the co-called Kozluk-Sason-Muş group. Based on Blanc's (1964) seminal book *Communal Dialects in Baghdad*, Anatolian Arabic is part of the *qəltu* dialects.[1]

Starting off with a geographical survey of Arabic dialects spoken in Turkey, Anatolian *qəltu*-dialects are generally argued to consist of four major groups (Jastrow 2006a):

---

1.   Blanc (1964) discusses the Arabic spoken in Baghdad, specifically in three religious communities, Muslim, Jewish, and Christian. Blanc noted that members of these communities spoke different dialects although they lived in the same town. Therefore, he classified the Jewish and Christian dialects as *qəltu* dialects and the Muslim dialect as *gilit* dialects, on the basis of the word *qultu* 'I said' of Classical Arabic.

(1)   i.   Mardin group
      ii.   Siirt group
      iii. Diyarbakır group
      iv.   Kozluk-Sason-Muş group

Our data comes from the variety spoken in the village of Kuzzang in the province of Mutki, Bitlis, and the village of Purşang, Batman.[2] The other languages spoken in the area are the official language of the country, Turkish, and Kurdish, Zazaki, and Armenian. Standard/Classical Arabic does not have any significant presence, other than in the religious sphere, and thus diglossia is not a critical factor. As expected in such situations, Sason speakers are typically multilingual. The map (from Jastrow 2006b) in Figure 1 marks the main geographical area where SA is spoken.
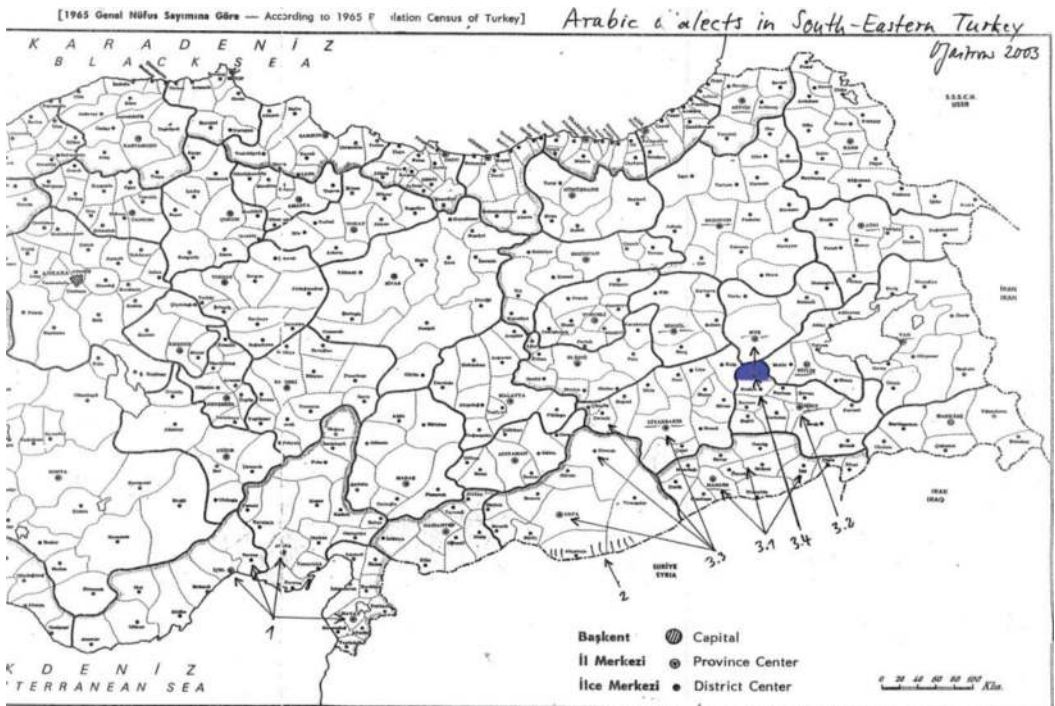
Figure 1. Sason Arabic and other Anatolian Arabic varieties

---

2.   For several reasons, e.g. the verbal modifications (Isaksson 2005: p. 187), we take the variety at hand to be different from the one discussed in Isaksson (2005) in the village of Xālile. For further information, see Akkuş (2016, to appear).

2.1   Sason morphology

Though it is obvious that SA is an Arabic variety, it has diverged from its Arabic relatives spoken in the Arab world. For example, it does not have as robust a root and pattern word formation system, though most of its morphology maintains its Arabic characteristics.[3] This section discusses a number of morphological properties of Sason Arabic, many of which we believe to be due to contact.

2.1.1   *Reduplication*
A type of reduplicative process due to contact with Turkish produces forms which are called *doublets with /m/* following Lewis's (1967: p. 237) account for Turkish. /m/ is added initially to words with initial vowels, as in (2a) or replaces the initial consonants in words with initial consonants, as in (2b). The new meaning is either that of *vagueness* or *et cetera*, along with the function of attention getting. It is reminiscent of the 'food shmood' type reduplication in English.

> (2)   a.   *asal   m-asal*
> honey m-honey
> 'honey or something like that'
> b.   *gerre   merre*
> noise m-noise
> 'noise or something like that'

2.1.2   *Degree in Adjectives*
Adjectives in SA are part of the noun phrase and follow the noun directly, agreeing with it in *gender*, *number*, and *definiteness*. In this respect, it is similar to what we find in most Arabic varieties. The category *degree*, on the other hand, is not an inflectional category. Instead, it has adopted the Turkish adverbs *daha* 'more' and *en* 'most' for comparative and superlative, respectively. The adverb *daha* 'more' precedes the adjectival constituent (3a), similarly the superlative adverb *en* 'most' comes before the adjective (3b).

> (3)   a.   *daha gbir*
> more big
> 'the bigger'
> b.   *en   gbir*
> most big
> 'the biggest'

---

3.   In this paper, we do not discuss the phonological properties of SA. The reader is referred to Akkuş (2016, to appear), Akkuş and Benmamoun (2016).

Lahdo (2009: p. 198) shows that the Tillo variety also uses the Turkish *en* 'most', realized as *an*, in superlative forms both in Arabic and Turkish words, as in (4).

(4) a. *an   atyap*
         most delicious
         'the most delicious'
     b. *an   yāqən*
         most close
         'the closest'

On the other hand, in the Tillo variety the comparative is formed through the *elative* (which functions both as comparative and superlative). In comparatives the preposition *mən* 'from' is used (Lahdo 2009: p. 162).

(5) *Təllo iyy   aṭyap mən əṢṭanbūl*
     Tillo be.PRES better than Istanbul
     'Tillo is better than Istanbul.'

This superlative form in Sason (and Tillo) is most likely due to contact with Turkish. This is not surprising since it is a complex construction in Arabic and has undergone changes in other Arabic dialects, though those changes do not seem to be due to contact.

2.1.3   *Compounding*

Sason Arabic seems to have borrowed the N + N compounding strategy from Turkish, where the compound linker is attached to the right-hand member (Kornfilt 1997; Göksel and Kerslake 2005). We generally do not find this pattern in other varieties of Arabic and it is most likely due to contact with Turkish. Note that these are not the so-called synthetic compounds, because the head noun does not carry any derivational morphology although the left-member of the compound is the theme of the head on the right.

(6) a. *lisa   mudur-i*
         high school director-CL
         'high school director'
     b. *qurs   oratman-i*
         course teacher-CL
         'course teacher'

In brief, there is no doubt that contact with Turkish and other neighboring languages has led to significant changes in the morphology and morphosyntax of Sason Arabic. For the rest of the paper, we will focus on other equally, if not more, important changes that we are confident are due to contact.

3.   Contact-induced morphosyntactic changes

Sason Arabic (as well as the Kozluk-Sason-Muş dialect group) manifests signifi-
cant contact-induced changes in the domain of syntax. In fact, it is probably the
Anatolian dialect with the most drastic changes due to contact. This section illus-
trates several syntactic constructions that could be attributed with a high degree of
certainty to change as a result of contact with the surrounding dominant languages,
primarily Turkish and Kurdish. These constructions include (in)definiteness mark-
ing, copular constructions, light verb constructions, and periphrastic causatives.

3.1   Indefiniteness marking

The marking of indefiniteness in Sason Arabic is a hallmark of the contact-induced
change, which in turn is connected with the loss of the definite marker that we
find in Classical Arabic and all the modern varieties spoken in the Arab world.
For instance, in Arabic dialects, indefinite NPs are unmarked or are preceded by
an independent indefinite particle, while an NP becomes definite by prefixing the
definite article *al-, əl-, il-* (Brustad 2000; Jastrow 2006a; Ryding 2005), e.g. *ʔaSiide*
'a poem', *l-ʔaSiide* 'the poem' in Lebanese Arabic.

   In Sason Arabic, on the other hand, the definite article has been preserved only
in a few frozen expressions, as illustrated in (7).

   (7)   *bı-**l**-xer   ci-to!*
         in-the-goodness came-2PL
         'Welcome!'

In most other contexts, the definiteness marker is entirely dropped.

   (8)   a.   *hatta   maytebe*
              until   school
              'until the school'
         b.   *mı   beyt*
              from   house
              'from the house'

Moreover, in addition to the loss of the definite article, Sason Arabic has also devel-
oped an indefiniteness marker that is enclitized to the noun, a pattern that is found
in Iranian and Turkic languages. Interestingly, the same change has been identified
in Uzbekistan Arabic as a result of its contact with Uzbek and Tajik (Jastrow 2006a),
both head final languages. Sason Arabic uses the enclitic *-ma* to mark the indef-
initeness of an NP. This indefinite element is unique to Sason group and is most
likely related to the Classical Arabic quantifier *-maa* 'something'.

(9)   a.   *kelp*
            'dog'
      b.   *kelp-ma*
           dog-a
           'a dog'

It seems to echo similar constructions we find in Kurdish and Turkish as illustrated in (10) and (11) respectively.

(10)   *derî*   > *derí-yek*   (Kurdish)
       the door   > a door

(11)   *kadın*   > ***bir*** *kadın*   (Turkish)
       the woman > a woman

The following examples show the marking of referentiality in Sason and its inter-action with word order (Akkuş, to appear).

(12)   a.   *naze masag-e atsūra*   non-referential   SVO
            'Naze caught a bird/birds' or 'Naze did bird-catching.'
       b.   *naze atsūra masag-əd-a*   definite, specific   SOV
            'Naze caught the bird.'
       c.   *naze masag-e atsūra-ma*   non-specific/indefinite[4]   SVO
            'Naze caught a bird.'

The unmarked word order in transitive sentences is SVO in Sason, and the position of the object may vary depending on its referential properties. The most salient reading for the bare noun *atsūra* in (12a) is an incorporation reading, in which the NP is number-neutral. It tends to express an activity interpretation (although the referential reading is also possible given the right context). In such cases, the default word order is SVO. In (12b), however, the same NP can only be interpreted as a definite expression since it has been moved to a pre-verbal position (thus, the SOV order) and more importantly an weak object pronoun is attached to the predicate to allow this reading. In (12c), the NP *atsūra* bears the postposition *-ma*, and it expresses an indefinite/nonspecific interpretation.

(13)   *naze atsūra-ma masag-əd-a*   specific/indefinite   SOV
       'Naze caught a certain bird' or 'A bird is such that Naze caught it.'

The example in (13) shows that what is being marked is not *definiteness*, but *specificity* (Heim 1982; von Heusinger and Kornfilt 2005), since the object pronoun can be combined with the indefinite article.

---

4.   The non-referential NP in (12a) can be distinguished from the indefinite in (12c) by the ability of the latter, but not the former, to pronominalize.

Crucially Turkish has the same four-way distinction in its marking of referentiality. It exhibits a morphosyntactic contrast between instances of the direct object with the case marker *-(y)I* and those without it. The accusative case suffix *-(y)I* indicates the specificity of its noun phrase (e.g. Enç 1991; von Heusinger and Kornfilt 2005).

Turkish does not have a definite article, but an indefinite article *bir*, related to the numeral *one*. The direct object can be realized as a bare noun without a case ending or as a noun (phrase) with the accusative case suffix *-(y)I* ((14) from von Heusinger and Kornfilt 2005).

(14)   a.   *(Ben) kitap oku-du-m.*   incorporated
            I   book read-PST-1SG
            'I was book-reading.'
       b.   *(Ben) kitab-ı   oku-du-m.*   definite
            I   book-ACC read-PST-1SG
            'I read the book.'
       c.   *(Ben) bir kitap oku-du-m.*   nonspecific, indefinite
            I   a   book read-PST-1SG
            'I read a book.'
       d.   *(Ben) bir kitab-ı   oku-du-m.*   indefinite specific
            I   a   book-ACC read-PST-1SG
            'I read a certain book.'

The change in the pattern is supported by the constructions that exhibit the *definiteness effect* (Milsark 1977). For instance, existential constructions disallow definite NPs: thus in English one can say *There is a bird on the roof*, but not *There is the bird on the roof*. Interestingly, in Sason bare nouns, without *-ma*, are not felicitous in this construction, and the absence of *-ma* renders the sentence ungrammatical, hence supporting the thesis that this property of Sason may be due to contact with Turkish.

(15)   a.   **ifi   atsūra fo fistox*
            there bird   on roof
            'There is the bird on the roof.'
       b.   *ifi   atsūra-ma fo fistox*
            there bird-a   on roof
            'There is a bird on the roof.'

Turkish (and also Kurdish) shows the same pattern, in the sense that the indefinite form is used in existential constructions.

(16)   *Çatı-da *(bir) kuş var.*   (Turkish)
       roof-LOC a   bird there is
       'There is a bird on the roof.'

The Sason and Turkish patterns are different from the pattern we find in Damascus (Syrian) Arabic:

(17) *kān fi   mara.*   (Damascus Arabic, Jastrow 2005: p. 135)
      was there a woman
      'There was a woman'

Another construction where this effect is observed is exclamatives such as 'what an XP…', which disallow definite NPs, as illustrated in English and its SA counterpart.

(18) a. What a/*the beautiful house it is!
      b. *beyt-*(ma) şıne   koys   ye*
          house-a   what beautiful COP.3
          'What a beautiful house it is!'

Putting aside some complexities related to these constructions and their theoretical implications, the important conclusion for the purposes of this paper is that the distribution of indefiniteness in Sason seems to pattern with what we find in Turkish (and Kurdish) and hence is likely due to contact.[5]

3.2   Light verb constructions

Light verb constructions are another domain where the influence of contact is clearly manifested. In surrounding languages, particularly Kurdish and Turkish, the form of light verbs is 'nominal + light verb', e.g. Kurdish *pacî kirin* (kiss do) 'to kiss', Turkish *rapor etmek* (report do) 'to report'.

    Light verb constructions in Sason are, not surprisingly, also formed with a nominal and the light verb *asi* 'to do'. The nominal part in Sason can be borrowed from Turkish as in (19d), or Kurdish as in (19c) or might be Arabic (19a–b). In Anatolian dialects, many expressions of this kind are found, not only with Turkish words, but also with Arabic words: *sawa talafōn* 'to phone', *sawa īšāra* 'to give a sign', *sawa mḥāfaza* 'to protect' (Versteegh 1997: p. 215).

(19) a. *meraq asi*
          wonder do
          'I wonder'

---

5.   Note that some Arabic varieties, such as Iraqi and Moroccan, have developed indefiniteness markers out of numerals. Our main point is that in Sason the semantic distribution of indefiniteness is similar to what we find in Turkish and Kurdish, which makes the case for contact as the most plausible trigger for the development of the marker.

    b.   *išāret asi*
        sign   do
        'I make a sign'
    c.   **ser**   *asi* < Kurdish
        watch do
        'I watch'
    d.   **qazan** *asi* < Turkish
        win   do
        'I win'

Versteegh (1997) argues that constructions with the verb *sawa* 'to do, make' in Anatolian Arabic is most likely 'a calque' of Turkish *etmek* 'to do'. The data provides support to Versteegh's argument. More importantly, Sason manifests the head final order for this light verb construction, undoubtedly due to contact.

3.3   Periphrastic causative

Sason Arabic resorts to periphrastic causative and applicative constructions rather than the root and pattern strategy found in other non-peripheral Arabic varieties. In this respect it is on a par with Kurdish, which uses the light verb *bıdın* 'give' to form the causative (20).

(20) *mı*   *piskilet*   *do*   *çekır-ın-e*
       1SG.OBL bicycle.NOM give.PART repair.PART-GERUND-OBL
       'I had the bicycle repaired (Lit: 'I gave the bicycle to repairing).'
                                        (Atlamaz 2012: p. 62)

Sason seems to deploy the same strategy for causative and applicative formation, as illustrated in (21). This could be as a result of its extensive contact with Kurdish.

(21) *adı-du*   *qattil*
       gave.3F.SG-him killing
       'She had him killed.' (Lit: 'She gave him to killing).

3.4   Negation and copula in Sason

Before discussing negation and the copula in Sason, let us briefly look at word order in sentences headed by verbs. As shown in (22a–d), SA display the VS and SV orders, which are also found in other Arabic dialects. Like those dialects, other orders are possible in certain pragmatic contexts (Akkuş 2014; Akkuş and Benmamoun 2016).

(22) a. *kemal qar-a   kitab-ad*   SVO
          K   read.PST-3M.SG book-PL
          'Kemal read books.'

      b. *qar-a   kemal kitab-ad*   VSO
          read.PST-3M.SG K   book-PL
          'Kemal read books.'

      c. *sabi-yad namo*   SV
          kid-PL   slept.3PL
          'The kids slept.'

      d. *namo   sabi-yad*   VS
          slept.3PL   kid-PL
          'The kids slept.'

As pointed out in Akkuş and Benmamoun (2016), the VS(O) order is generally found in relative clauses as in (23a) and embedded clauses as in (23b).

(23) a. *ınt   kitab le   i-habb   cihan tı-qri*
          2M.SG   book   that 3M.SG-love   Cihan 2M.SG-read
          'You read the book that Cihan likes.'

      b. *ma-sıma-tu   le   jo   zġar*
          NEG-heard-1SG   that came.3PL   children
          'I didn't hear that the children came.'

In this respect, SA has maintained the syntactic characteristics typical of its relatives spoken in the Arab world. We turn next to negation and copula constructions where word order has undergone significant changes.

3.4.1   *Negation*

Like most Arabic dialects, the main sentential negative in Sason is realized as the proclitic *maa* or, depending on tense, a variant of *maa* (such as *mo/mi*):

(24) *maa adaš-tu   tunes.   (ma:daštu)*
      NEG saw-1SG   anyone
      'I didn't see anyone.'

Again, like other Arabic dialects, particularly those spoken in Iraq and the Gulf region, the negative in imperative constructions is *laa*.

(25) *laa   tamel.*
      NEG work.2M.SG
      'Don't work.'

Table 1 provides the paradigm of negation in sentences with verbal predicates.

Table 1. Negative markers in Sason

Tense	Negative particle
Present/Future (non-past)	*mo-/mı-/mi-*
Past	*maa*
Imperative	*laa*

3.4.2   Copular constructions

One well known characteristic of most varieties of Arabic, including Classical Arabic, is the absence of an overt copula in non-generic present tense sentences (Benmamoun 2000; Aoun et al. 2010; Benmamoun et al. 2014). Interestingly, SA deviates from this pattern in that it has evolved a copula in those same contexts where it is absent in other Arabic varieties spoken in the Arab world. Consider the sentences in (26):

(26) a. *ab-i   nihane *(ye)*
          father-my here   COP.3SG
          'My father is here.'
      b. *nihane   kıntu*
          nihane   be.1SG.PRES
          'I am here.'

In (26a, b) the copula markers *ye* and *kıntu* are required and they occur in both main or matrix sentences as in (26) and in dependent or embedded clauses as in (27).

(27) *mo-saddex   le   Naze raxu-e ye*
      NEG-1SG.believe that Naze sick-F   COP.3.SG
      'I don't believe that Naze is sick.'

The full paradigms of the copula found in the present and past tenses are given in Tables 2 and 3 respectively (from Akkuş and Benmamoun 2016).

As pointed out in Akkuş and Benmamoun (2016), the present tense copula paradigm seems to have drawn from the pronominal paradigm (third person) and the past tense copula (first and second persons).

The head-final order in (27) is most likely the result of the head final neighboring languages which have copulas in the present tense and which are placed after the predicate (from Grigore 2007).[6]

---

6. Turoyo is a Semitic (Neo-Aramaic) language.

Table 2. Present tense Copula

Person	Number	Gender	Auxiliary
1	Singular	M/F	*kıntu* 'I am'
2	SG	M	*kınt*
2	SG	F	*kınte*
3	SG	M	*ye*
3	SG	F	*ye*
1	Plural	M/F	*kınna*
2	PL	M/F	*kınto*
3	PL	M/F	*nen*

Table 3. Past tense Copula

Person	Number	Gender	Auxiliary
1	Singular	M/F	*kıntu* 'I was'
2	SG	M	*kınt*
2	SG	F	*kınte*
3	SG	M	*kan*
3	SG	F	*kane*
1	Plural	M/F	*kınna*
2	PL	M/F	*kınto*
3	PL	M/F	*kano*

(28) a. *bave   minŞivan-e*   (Kurdish)
          father   shepherd-3SG
      b. *babam   çoban-dır*   (Turkish)
      c. *babi   rəʕyo-yo*   (Turoyo)
          'My father is a herder'

In negative sentences, NEG (and the copula if there is one) follows the predicate, as the sentences in (29) show.

(29) a. *hasta   değil-ler*   (Turkish)
          sick   NEG-COP.3PL
          'They are not sick'
      b. *kemal   xwandekar nin-a*   (Kurdish)
          Kemal   student   NEG-3SG
      c. *cinya   niwaš ni-yo*   (Zazaki)
          child   sick   NEG-3SG

Our contention, discussed in Akkuş and Benmamoun (2016), is that the main driver for the development of the head final copula was most likely contact with

Kurdish. The copula could also have evolved under contact with Turoyo with the head final order being influenced by contact with Kurdish. In other words, it may well be the case that while this is contact-induced change, more than one contact language may have played a role. Needless to say, this important question requires further research.

4. Conclusion

In this paper, we discussed various contact-induced morphosyntactic and syntactic changes in Sason Arabic, spoken in eastern Turkey. The changes are many but we have focused mainly on indefiniteness, light verb constructions, causatives, and negative copula sentences. In all those constructions, Sason Arabic exhibited patterns that are more in line with the languages it is in contact with than with its Arabic relatives, which makes it highly plausible that the patterns, particularly the word order patterns, are due to contact rather than language internal change. At this point in our research on Sason Arabic we cannot think of any compelling Sason-internal syntactic reasons that would have given rise to the above patterns. We feel confident at this point to attribute them mainly to contact.

Regarding the mechanism involved and the limits of syntactic change, Ratcliffe (2005: p. 141) asks the following significant questions:

i.   Are there any limitations on what aspects of language can change due to contact with another language?
ii.   How are reversals of normal word order patterns implemented in the course of a change of type?

Ratcliffe (2005) suggests, based on his analysis of Bukhara Arabic, that this dialect of Arabic seems to fit a pattern where morphosyntactic structure is more vulnerable to change under pressure from neighboring languages, particularly Uzbek (an SOV language). This accounts, for example, for the SOV clause structure and the head final nature of relative clauses. However, Ratcliffe and, as far as we can determine, others have not shown how this change in Arabic varieties actually has unfolded. In general, the standard and sensible approach attributes the change to bilingualism/multilingualism, but how syntactic change has actually developed is left open. We recognize that this is a difficult question for most spoken Arabic varieties, peripheral or non-peripheral, which are usually not written and are not used as vehicles for literacy. We speculate that one path for word order change, though by no means the only one, that Sason illustrates is through structures that do not have parallels in the language. Sason Arabic seems to provide evidence that such structures, such as copula constructions, are a 'Trojan horse' to catalyse change in

syntactic patterns, particularly word order. If this is correct, we expect the trend to extend to other constructions such as sentences with verbal predicates, relative clauses and other complex syntactic constructions. Eventually, Sason may end up looking more like Bukhara Arabic, with a typically Arabic morphological structure but a head final syntactic profile.

List of abbreviations

ACC	accusative	NOM	nominative
CL	compound linker	OBL	oblique
COP	copula	PART	?
F	feminine	PST	past
GERUND	gerundive	PL	plural
LOC	locative	PRS	present
M	masculine	SG	singular
NEG	negation		

References

Akkuş, F. (2014). *The Functional Categories and Phrase Structure of Sason Arabic*. MA thesis, Boğaziçi University.

Akkuş, F. (2016). The Arabic Dialect of Mutki-Sason Areas. In G. Grigore and G. Biţună (eds.), *Arabic Varieties: Far and Wide. Proceedings of the 11th Conference of AIDA*. Bucharest: Editura Universităţii din Bucureşti. 29–41.

Akkuş, F. (To appear). Sason Arabic. R. de Jong (ed.). *Encyclopedia of Arabic Language and Linguistics Online*. Leiden: Brill.

Akkuş, F. and E. Benmamoun (2016). Clause Structure in Contact Contexts: the Case of Sason Arabic. *Perspectives on Arabic Linguistics* XXVIII. Y. Haddad and E. Potsdam (eds). Amsterdam: John Benjamins. 153–172.

Aoun, J., E. Benmamoun and L. Choueiri (2010). *The Syntax of Arabic*. Cambridge: Cambridge University Press.

Atlamaz, Ü. (2012). *Ergative as Accusative Case: Evidence from Adıyaman Kurmanji*. MA thesis. Boğaziçi University.

Benmamoun, E. (1992). *Inflectional and Functional Morphology: Problems of Projection, Representation and Derivation*. Unpublished Ph.D. dissertation, University of Southern California, Los Angeles.

Benmamoun, E. (2000). *The Feature Structure of Functional Categories: A Comparative Study of Arabic Dialects*. New York: Oxford University Press.

Benmamoun, E. et al. (2014). The Location of Sentential Negation in Arabic Varieties. *Brill's Annual of Afroasiatic Languages and Linguistics* 5: 83–116.

Blanc, H. (1964). *Communal Dialects in Baghdad*. Cambridge, MA: Harvard University Press (Harvard Middle Eastern Monographs X).

Brustad, K. (2000). *The syntax of spoken Arabic*. Washington, DC: Georgetown University Press.

Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry* 22: 1–25.

Göksel, A., and C. Kerslake (2005). *Turkish: A Comprehensive Grammar*. London and New York: Routledge. https://doi.org/10.4324/9780203340769

Grigore, G. (2007). L'énoncé non verbal dans l'arabe parlé à Mardin. *Romano-Arabica* N: 6–7.

Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. Ph.D. dissertation, University of Massachusetts, Amherst.

Isaksson, B. (2005). New Linguistic Data from the Sason Area in Anatolia. In Á. Csató, B. Isaksson, and C. Jahani (eds.), *Linguistic Convergence and Areal Divergence: Case Studies from Iranian, Semitic and Turkic*. Routledge. New York. 181–190.

Jastrow, O. (1978). *Die mesopotamisch-arabischen qəltu-Dialekte. Band I: Phonologie und Morphologie*. Wiesbaden: Steiner.

Jastrow, O. (2005). Uzbekistan Arabic: A Language Created by Semitic-Iranian-Turkic Linguistic Convergence. In Á. Csató, B. Isaksson, and C. Jahani (eds.), *Linguistic Convergence and Areal Divergence: Case Studies from Iranian, Semitic and Turkic*. Routledge. New York. 133–139.

Jastrow, O. (2006a). Anatolian Arabic. *Encyclopedia of Arabic Language and Linguistics*. Kees Versteegh (ed.). 1. 86–96.

Jastrow, O. (2006b). Arabic dialects in Turkey – towards a comparative typology. *Workshop on Turkish Dialects Orient Institute*, Türk Dilleri Araştırmaları. Istanbul. 153–164.

Kornfilt, J. (1997). *Turkish grammar*. NY: Routledge.

Lahdo, A. (2009). *The Arabic Dialect of Tillo in the Region of Siirt (South-eastern Turkey)*. Acta Universitatis Upsaliensis. Studia Semitica Upsaliensia 26. Uppsala.

Lewis, G. (1967). *Turkish Grammar*. London: Oxford University Press.

Milsark, G. (1974). *Existential Sentences in English*. Ph.D. dissertation, MIT.

Milsark, G. (1977). Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3: 1–29.

Ratcliffe, R. R. (2005). Bukhara Arabic: a metatypized dialect of Arabic in Central Asia. In Á. Csató, B. Isaksson, and C. Jahani (eds.), *Linguistic convergence and diffusion. Case studies from Iranian, Semitic, and Turkic*. London: Routledge. 141–159.

Ryding, K. (2005). *A Reference Grammar of Modern Standard Arabic*. Cambridge University Press. https://doi.org/10.1017/CBO9780511486975

Talay, S. (2001). Der Arabische Dialekt von Hasköy (Der Khas), Ostanatolien I. Grammatische Skizze. *Zeitschrift für arabische Linguistik* 40: 71–89.

Versteegh, K. (1997). *The Arabic Language*. Columbia University Press.

von Heusinger, K. and J. Kornfilt (2005). The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic Languages* 9: 3–44.

Arabic-Berber-Songhay contact and the grammaticalisation of 'thing'

Lameen Souag

The development of double negation in Arabic has attracted increasing attention in recent years. The striking parallels between negation in Berber and North African Arabic invite an explanation in contact terms, and such explanations have indeed been debated. However, in addition to their use in postverbal negation, reflexes of *šayʔ* have several functions not directly related to negation, notably indefinite quantification and polar question marking. The marking of these functions, too, shows striking Arabic-Berber parallels generally neglected in discussions of the phenomenon. Taking these into account produces a more complete picture of contact influence, and provides clues to the relative chronology of these developments. In some cases, non-Arabic varieties are found to preserve usages obsolete in present-day regional Arabic dialects.

Keywords: language contact, contact grammaticalisation, calquing, negation, interrogation, quantification, Arabic, Berber, Songhay

1. Introduction

Over more than a millennium of intense contact, Berber and Arabic have profoundly influenced one another in North Africa. In some cases, Berber varieties preserve forms and constructions long since abandoned by their present-day Arabic-speaking neighbours (Souag 2009; Souag 2015a). Berber evidence is nevertheless often neglected in attempts to unravel the history of Arabic dialects. In the case of double negation, while many discussions ignore Berber entirely (e.g. Diem 2014; Wilmsen 2014), the importance of Arabic influence on Berber in the development of double negation has already been highlighted (Lucas 2010: pp. 50–64; Lucas 2013). In both Arabic and Berber, in precisely the same contexts, a word originally meaning 'thing' has developed into a postverbal negative particle, apparently following the familiar cycle of count noun (/_NEG) > emphatic negator (/_NEG) > NEG2 (/_NEG) > NEG2 outlined by Jespersen (1917); notwithstanding the scepticism

of Brugnatelli (2014), this seems unlikely to be coincidental. Contact grammaticalisation of this kind is widely attested worldwide and is reasonably well-understood (Heine & Kuteva 2005); it is driven primarily by the calquing of polyfunctionality.

However, negation is far from being the only function developed by Arabic *šayʔ* 'thing': as already highlighted by Obler (1990), this morpheme's descendants have a wide variety of functions in Arabic dialects, of which the most widespread besides negation include indefinite quantification and polar interrogation. If the development of 'thing' in Berber has been influenced by Arabic, one might expect it to have developed the same functions in Berber as well. This expectation is in fact borne out by the data. However, discussion of the role of contact in grammaticalisation from 'thing' in the Arabic and Berber literature focuses almost exclusively on negation. Kossmann (2013: pp. 305–306) briefly discusses the possibility of Arabic influence on the development of polar interrogation marking in Berber, but gives little space to the Arabic dialectological evidence for this. In Korandje – a heavily Berber- and Arabic-influenced Northern Songhay variety spoken in the Algerian Sahara – parallel developments are also observable, and have not yet been discussed anywhere. In this paper, evidence will be presented for the claim that the grammaticalisation of "thing" within Arabic has influenced Berber, and thence Northern Songhay, not only in the domain of negation but also in the domains of quantification and interrogation.

2.   The pre-contact functions of Arabic *šayʔ* and Berber **kăra*

Prior to intensive contact between Arabic and Berber, each seems to have developed – presumably independently, or at most mediated by more remote areal factors – a word covering the senses 'thing', 'something', and, under negation, 'nothing'. The evidence for this is direct in the case of Arabic, and indirect but strong in the case of Berber.

2.1   Arabic

As Al-Jallad (2014) points out, Arabic *šayʔ* 'thing' appears to be historically a deverbal noun from the root *šyʔ* 'to want, to will', attested in earlier Safaitic with the more specific meaning 'to experience lack or want'. The implied development of 'need' > 'thing' precisely parallels the more recent history of widespread dialectal Arabic *ħājah*. By the beginning of the Islamic era, it had acquired a further function as an indefinite pronoun 'anything'; both usages are well-attested in the Qurʔān and would become standard in Classical Arabic:

(1) *ʔin t-ubd-ū   šayʔ-a-n   ʔaw t-uxf-ū-hu…*
    if 2IPFV-reveal-2M.PL thing-ACC-INDF or 2IPFV-conceal-2M.PL
    'Whether you reveal anything or conceal it …'   (Qurʔān 33: p. 54)

As such, the indefinite accusative ending allowed it to form adverbial 'at all':

(2) *fa-l-yawm-a   lā   t-uḍlam-u   nafs-u-n*
    for-DEF-day-ACC NEG 3F.SG.IPFV-wrong.PASS-DECL soul-NOM-INDF
    šayʔ-a-n
    thing-ACC-INDF
    'This day no soul will be wronged in any way / at all.'   (Qur'ān 36: p. 54)

Of course, written sources represent only a small part of the dialectal diversity of early Arabic; it is possible that *šayʔ* had already acquired more functions in some dialects. However, these can at any rate be taken as the minimal set of functions for *šayʔ* when Arabic first came into contact with Berber during the early Islamic era.

2.2   Berber

No relevant documentary evidence on Berber exists for the early Islamic period, so it is necessary to take a more indirect approach to the reconstruction of the functions of *ḱăra. This etymon is found everywhere in Berber except in Tuareg, whose *hărăt* 'thing, something, a little of' cannot be regularly related to *ḱăra despite its obvious phonetic and semantic similarity. The subclassification of Berber remains problematic, but Zenaga (in Mauritania) stands out as a clear outlier (Basset 1952; Kossmann 1999), along with its recently discovered close relative Tetserrét (Lux 2013); only slightly less divergent are the Libyan varieties of Ghadames and Awjila. Moreover, Zenaga and Awjila are both spoken at the extreme peripheries of Berber, making it relatively unlikely that contact should affect them both in identical ways. Zenaga is particularly valuable in that it also shows unusually little Arabic influence. Any function of *ḱăra shared by Zenaga with at least one of the Libyan varieties mentioned is thus effectively certain to be proto-Berber. All four of them share the nominal sense 'thing' for this etymon (whose various senses will for convenience be glossed below as KRA); this sense may thus be taken as original:

(3) Zenaga: *tôbbəl   tə-skər…   əgălləǰʰ   ən   əď   karəh*
    (Mauritania) servant.F 3F.SG-make.PFV… many   GEN PL   KRA
    'The maidservant made … many things.'   (Nicolas 1953: p. 73)

(4) Tetserrét: *karad wad   an   tə-gas-ət*
    (Niger)   KRA   DEM.M CENTRIF 2-find.PFV-2SG
    'the thing that you have found'   (Khamed Attayoub 2001: p. 143)

(5) Ghadames: *i-ssə-mūd*   *kara-y-o*
    (W Libya)   3M.SG-CAUS-pray KRA-EP-DEM
    'This thing/act makes prayer valid.'   (Lanfry 1973: p. 159)

(6) Awjila: *kéra-y-âya*   *ddîwa*
    (E Libya)   KRA-EP-DEM   what
    'What is this thing?'   (van Putten 2014: p. 263)

At least three out of these four varieties also share the more grammaticalised sense of 'something', or with negation 'nothing', with most of the more central branches of Berber:

'something':

(7) Zenaga: *y-ukf-iʔh*   *kāräh*
              3M.SG-give.PFV-1SG.DAT KRA
    'He gave me something.'   (Taine-Cheikh 2008: p. 299)

(8) Tetserrét: *har*   *i-ḍaš*   *karad*
               until 3M.SG-touch KRA
    'Until he touched something …'   (Lux 2013: p. 556)

(9) Ghadames: *awādəm*   *i-ttă-tħăššăm*   *d-i-ftăk*
              human   3M.SG-IPFV-feel.shame   IRR-3M.SG-beg
    kara   *ħarmān*
    KRA   deprivation
    'A person is ashamed to beg for something out of deprivation.'
                                                           (Lanfry 1968: p. 4)

'nothing' (under negation):

(10) Zenaga: *wär-iʔh*   *y-ukfi*   *kārāh*
               NEG-1SG.DAT   3M.SG-give.PFV   KRA
     'He gave me nothing.'   (Taine-Cheikh 2008: p. 299)

(11) Tetserrét: *iwwat* *ešli*   *ad*   *wur* *n-ila-t*   *karad*
                 one.F   woman   DEM   NEG   PTCP-have.PFV-3SG.ACC   KRA
     'a woman who had nothing'   (Khamed Attayoub 2001: p. 135)

(12) Awjila: *wur*   *ġâr-i*   *kéra*
                NEG   at-1SG   KRA
     'I have nothing.'   (van Putten 2014: p. 263)

The extension from the noun 'thing' to the indefinite pronoun 'something/nothing' can thus plausibly be considered to have taken place before the breakup of proto-Berber. These two senses are, in fact, the only ones well-attested in both of the two peripheries, and thus the only ones that can be confidently reconstructed for the proto-Berber stage. While 'thing' is commonly expressed in modern northern

Berber varieties by Arabic loans, a number of varieties from Tashelhiyt in southern Morocco through to Chaouia in eastern Algeria use the Latin loan *ta-ɣawsa < causa (Kossmann 2013: p. 71). The grammaticalisation from 'thing' to 'something' in Berber must therefore have taken place before Latin ceased to be influential in North Africa, and hence most probably before Arabic contact became significant. At the time of first Arabic-Berber contact, the three main senses of Arabic šayʔ – 'thing', 'something', and 'nothing' – could therefore, at least in some areas, be translated by Berber *ḱăra (though the sense 'thing' may already have been lost in other areas).

3.   From indefinite pronoun to indefinite quantifier

The leap from indefinite pronoun to indefinite quantifier is fairly natural, but by no means inevitable, as attested by its absence in many varieties. It is widely attested in Arabic and extremely widespread in Berber; however, its distribution within Arabic includes dialects that can hardly be taken to have undergone influence from Berber, whereas the reverse is not true.

3.1   Arabic

The earliest examples of the use of šē/šī as an indefinite quantifier directly preceding the noun documented by Diem (2014: pp. 103–105) date to the 12th and 13th centuries, e.g.:

> (14)   Andalusi: *lāḥ*   *l-ak*   *šī*   *ʿimāra*
>   appear.3M.SG.PFV   to-2SG   thing   sign
>   'There appeared some sign to you.'   (Al-Shushtarī, d. 1269)

The early examples cited by Diem include Andalusi poetry and Geniza letters, variously probably or certainly from the Maghreb; there are no unambiguously Egyptian examples. All involve inanimates, in both realis and irrealis contexts. The construction continues to be widespread in modern Maghrebi Arabic; it is described in detail for Moroccan Arabic by Caubet (1983). Within the Maghreb, it is best attested in Morocco and western Algeria (Madouni-La Peyre 2003: p. 251), areas where a Berber substrate is prominent. In the eastern Maghreb, it is less widely reported; however, it is attested in Malta, eg *xi mkien* (< x > = š) 'somewhere' (Haspelmath & Caruana 1996: p. 215), and even in Benghazi in eastern Libya: *šī yowm* 'some day' (Benkato 2014: p. 88). In both the latter regions, Berber substratum influence hardly exceeds a few lexical items.

This use of *šē/šī* is, however, also attested in the Middle East. In Levantine Arabic, the use of *šē/šī* as an indefinite quantifier directly preceding the noun is well-established, and is a well-known feature of the regional koines. Among more local dialects, it is specifically reported (as *ši*) for Kfar `Abida in northern Lebanon (Feghali 1919: p. 279), (as *šī*) for Damascus (Cowell 1964: p. 467), (as *šē*) for Mharde in east-central Syria (Yoseph 2012: p. 54), (as *šī*) for the central Syrian oasis of Soukhne (Behnstedt 1994: p. 123). It is, moreover, used (as *ši*) in Cypriot Maronite Arabic (Borg 2004: p. 303), whose separation from the rest of Levantine is rather early. Further east, however, this usage is absent; in Iraqi Arabic koine, it is rather the more classical construction *šī min* which is used (Woodhead & Beene 1967: p. 254), while for Mardin, in Turkey just across from northeastern Syria, no such usage of *šī* is reported (Grigoire 2009: p. 241). A similar development with a different word order may be attested in Tillo in southeastern Turkey, as suggested by the following example:

(15) *ğīb=li*   *kəbrīt šī   mən ạwnak!*
      bring=1SG.DAT match some from there
      'Give me some matches over there!'   (Lahdo 2009: p. 229)

In south Arabia, *šē/šī* is not attested as an indefinite quantifier. However, it is attested throughout much of the region in a closely related function: as an irrealis existential, typically followed directly by the noun it predicates. For the Hadramawt dialect, Landberg (1901: p. 628) notes that "Il est à remarquer que شي n'est ainsi employé [comme "il y a"] qu'après une conjonction et dans une proposition interrogative ou négative" [It is to be noted that *šy* is used in this way (as "there is") only after a conjunction and in an interrogative or negative proposition]; this use of *šī* is widely attested in Yemen (Behnstedt 1996: p. 690). For Dhofar, likewise, existential *šē* is documented by Davey (2013: p. 170), and the examples given are consistently irrealis. As far north as Al-Baha in the southern Hijaz, *mā š(ī)* is used as the negative existential (Nadwi 1968: p. 138–139). In parts of eastern Oman, we even find this use of *šī* combined with suffixed *šī* negation: *šīši* 'there is not' (Brockett 1985: p. 91).

Given the Yemeni and Syrian data, it appears impossible to explain the Arabic development through contact with Berber. Its distribution within North Africa proper – specifically, its concentration in the northwest – is thus to be explained not by the stronger influence of Berber in the northwest, but rather by indefinite quantifier *šī* being a pre-Hilalian archaism within North African Arabic, brought from the east by some of the earliest Arabic-speaking immigrants and later receding under the influence of more recent arrivals using other strategies. Its attestation in Maltese, which has been isolated from other Arabic dialects since before the 11th century arrival of the Banu Hilal, strengthens this interpretation.

3.2 Berber

In Western Berber, the development from indefinite pronoun to indefinite quantifier remains relatively restricted. There, KRA is used only with mass nouns, meaning 'some, a little' – followed by a genitive preposition in Zenaga, but a locative one in Tetserrét:

(16) Zenaga: *kāräʔ-n   iʔžž*
      (Mauritania) KRA-GEN   milk
      'some milk, a little milk'   (Taine-Cheikh 2008: p. 299)

(17) Tetserrét: *karad gud tannəttan əd   təferdi-s*
      (Niger)   KRA   in   honey   and wax-3SG.GEN
      'some honey and its wax'   (Khamed Attayoub 2001: p. 137)

On this basis, it might be suggested that KRA was already used as an indefinite quantifier for mass nouns (but not count nouns) in proto-Berber. However, little confidence can be placed in this suggestion. The difference in constructions between Zenaga and Tetserrét suggests independent development (unless, indeed, the Tetserret construction – taken from an oral commentary on a textbook of Islamic law – is simply a calque from Arabic). That being the case, either or both could equally be recent developments.

    In most of Berber, however (including almost all Moroccan and Algerian varieties), KRA is used, followed by a genitive, as an indefinite quantifier for both count and mass nouns, e.g.:

(18) Kabyle: *kra   n   waman   / kra   n   tebratin*
      (Algeria)   KRA GEN water.ANN   KRA GEN letters.ANN
      'some water' / 'some letters'[1]

This usage is attested relatively early, as in the Zouagha manuscript (from a northwestern Libyan port where Berber is no longer spoken):

(19) Zouagha: <ʔđ   wys   ʕylʔ   šrʔ   ʕnwđrym   đrwsyⱦ>
      (W Libya)   *əđ   wis   i-la   šra   ən=wəḏrim   əđrus-iṯ*
                  COP REL.M.SG 3M.SG-have KRA GEN=money.ANN few-3SG.M
      '[A poor man] is he who has some money, (but) little.'
                                          (Calassanti-Motylinski 1905: p. 74)

It is even attested (with a count noun) in Awjila:

---

1. http://tatoeba.org/fra/sentences/show/2294432; http://tatoeba.org/fra/sentences/show/2810426.

(20) Awjila: *u=kan ət-ɣəlli-m   kəra   n=əlḥažət   t-ṣuṣm-im   ka*
      (E Libya)   and=if 2-want-2PL   KRA   GEN=thing   2-be.quiet-2PL   NEG
      'And if you pl. want something, don't be quiet.'

<div align="right">(Souag & van Putten 2015: p. 42)</div>

This suggests a relatively early date for this innovation – at a period after Zenaga and Ghadames had already separated from the mainstream of Berber, but well before the continuity of Berber across North Africa proper was broken up, and hence presumably before the 11th century. This is compatible with an explanation in terms of Arabic influence: we have seen above that the same *terminus ante quem* applies to this development within North African Arabic, and that its distribution makes it unlikely to have spread from Berber to Arabic. Independent development cannot be excluded completely without a *terminus post quem*; however, the northerly distribution of this innovation matches well with the distribution of the greatest levels of Arabic influence within Berber.

4.   From indefinite pronoun to indefinite adverb

In Classical Arabic, as noted, the adverbial usage of *šayʔ-an* is well-attested. Diem (2014: 13–21) gives a close study of its usage in the earliest Classical texts. In this period, it is found sentence-finally with positive declarative clauses in the sense 'a bit, somewhat' as well as with interrogatives and negatives in the sense 'at all, whatsoever'. The predicate is always gradable. If the verb is transitive, it must have an explicit object, normally definite. Examples (Diem 2014: p. 17–18) include:

(21)   *fa-danaw-tu   min-hu   šayʔ-a-n*
        so-approach-1SG.PFV   from-3M.SG   thing-ACC-INDEF
        'I approached him a bit.'   (*Musnad Ibn Ḥanbal* no. 15339)

(22)   *fa-hal ðālika   nāfiʕ-u-hā   šayʔ-a-n?*
        so-Q   that.M.SG   benefit.PTCP-NOM-3F.SG   thing-ACC-INDEF
        'Does this profit her a bit?' (rather: '…at all?') (*Musnad Ibn Ḥanbal* no. 15358)

(23)   *mā naqam-nā   min-hu   šayʔ-a-n*
        NEG revenge-1PL.PFV   from-3M.SG   thing-ACC-INDEF
        'We took no revenge whatsoever on him.'   (Ṣaḥīḥ Muslim no. 3407)

In positive contexts, this usage was replaced in the post-classical period by reflexes of the diminutive *šuwayʔ* (Diem 2014: p. 64). In negative contexts, it remained in usage much longer, behaving as little more than an emphatic negative particle and tending to cliticize to the verb; e.g. (Diem 2014: p. 39):

(24) Andalusi: *la   ta-n-qasám-ši*
                  NEG   2M.SG-PASS-divide-thing
      'you are not divided at all'   (al-Shushtarī, d. 1269)

This emphatic negative usage is obsolete. By the 17th century, it had already been reduced to an unmarked postverbal negator in Egyptian Arabic (Diem 2014: pp. 47–56), and no dialect of which I am aware continues to limit it to emphatic negation: either it marks negation in general, a phenomenon already amply discussed elsewhere, or it has no surviving negative adverbial function (although in some cases a full reflex *ši/šay* may mark emphatic negation in contrast with a coeval reduced reflex *-š*).

In interrogative contexts, however, reflexes of *šayʔ* are quite widely used to mark polar interrogation, including in the Levant, Cairo, Malta, Libya, Tunisia, Algeria, Morocco, and (arguably) Yemen (Singer 1958: pp. 81–91). Singer derives this function directly from the sense 'something'. In Moroccan Arabic, however, polar interrogative *ši* occurs only with gradable predicates (Caubet 1993, vol. II: p. 76), whereas no dialect seems to restrict it to antipassives; it thus appears more probable that this usage derives from the adverbial usage of the indefinite pronoun discussed earlier. In some cases, these reflexes are still straightforwardly adverbial: standalone adverbial reflexes of *šayʔ-a-n* remain in use as polar interrogation markers in Levantine dialects, e.g. Syrian:

(25) *ṭʔūmt-ī   ʔəž-ət   mən ʕand əl-kawwa   ši?*
      suit-POSS.1SG   come-3F.SG.PFV from at   DEF-cleaner Q?
      'Have my suits come back from the cleaners?'   (Cowell 1964: p. 378)

In North Africa, this usage is still possible in Eastern Libyan Arabic:

(26) *šif-t   aḥmad   amis   ši*
      see-2M.SG.PFV Ahmad yesterday Q
      'Did you see Ahmad yesterday?'   (Owens 1984: p. 102)

However, as in the negative usage, it shows a strong tendency to be reduced to a postverbal clitic; even in Eastern Libyan Arabic, it can also appear as a postverbal clitic *-š*.

5.   From indefinite adverb to polar interrogative marker

We have seen that the use of Classical Arabic *šayʔ-a-n* in interrogatives in the sense of 'at all, in any way' survives adverbially among Arabic dialects as a polar interrogative marker. This adverb has in turn frequently been reduced to a postverbal clitic,

both in Arabic and among the more strongly Arabic-influenced easterly Berber varieties.

5.1   Arabic

In most North African examples, interrogative *šayʔ* appears immediately after the verb, e.g.:

(27)   Algiers:[2] *<djaoueb   chi>*
                   jawǝb   ši
                   answer.3M.SG.PFV   Q
     'Did he answer?'   (Delaporte 1845: p. 82)

or:

(28)   Tunis: *t-uqʕód=ši   úlla t-ímši*
                 2SG.IPFV-stay=Q   or   2SG.IPFV=walk
     'Will you stay, or leave?'   (Stumme 1896: p. 143)

It can also directly follow a non-verbal predicator in some contexts, as in:

(29)   Algiers: *<And-ek=chi bezzaf menn-ou?>*
                 ʕand-ǝk=ši   bǝzzaf mǝnn-u?
                 at-2SG=Q   a.lot   from-3M.SG
     'Do you have a lot of it?'   (Cotelle 1847: p. 94)

Moroccan Arabic also shows a clause-initial interrogative particle *waš* (Caubet 1993, vol. II: p. 86). This form's precise relationship to interrogative *ši* is debatable; in at least some dialects, it is distinct from *aš* 'what?' Wilmsen (2014: p. 97) suggests a derivation from 3msg *huwa* (itself used phrase-initially to mark polar questions in several Arabic dialects) + interrogative *-š*; this is semantically plausible, but the loss of the *h* is difficult to explain.

5.2   Berber

In Berber, similar uses of **kǎra* are much rarer than in Arabic, but are nevertheless well-attested in a number of varieties, most of them particularly Arabized. The early 19th century Kabyle dialect of Bejaia described by Brosselard et al. (1844) seems systematically to mark polar questions with a morpheme *kra*, a usage

---

2.   The 19th century usage exemplified in these examples is obsolete in modern Algiers.

unknown in modern Kabyle (note that the town of Bejaia proper was predominantly Arabic-speaking at the time):

(30) Kabyle of Bejaia: <*Te-zemmer-eth   kra   a-yi-t-sahhah-eth*
      (Algeria)   2-can.IPFV-2SG   KRA   IRR-1SG.ACC-2-correct-2SG
      ouayi?>
      this.M.SG
      'Can you correct this for me?'   (Brosselard 1844: p. 21)

More frequently, we find a polar question marker *š*, optional or restricted to certain contexts, in varieties which also use *š* as NEG2 (cf. briefly Kossmann 2013: p. 306):

(31) Beni Snous: *i-llä-š*   *ġr-es   ụéġṛům*
      (W Algeria)   3M.SG-be.PFV-Q   at-3SG   bread.ANN
      'Do you have any bread?'   (Destaing 1907: p. 132)

(32) Chaoui of Batna: <*adz-i-ouett-ech*>
      (E Algeria)   *aḏ-i-wətt-əš*
                        IRR-3M.SG-hit-Q
      'Will he hit?'   (Torchon 1871: p. 82)

(33) Tamezret: *i-qam-ak-š*
      (S Tunisia)   3M.SG-lift.PFV-2M.SG.ACC-Q
      'Has he lifted you up?'   (Ben Mamou 2005; Kossmann 2013: p. 306)

(34) Zraoua: *t-ufí-š-š*   *ħáža?*
      (S Tunisia)   2-find.PFV-2SG-Q   something
      'Did you find anything?'   (author's field notes)

(35) Sened: *t-esen-et-ch*   *manet i-nr'a?*
      (Tunisia)   2-know.PFV-2SG-Q   who   3M.SG-kill.PFV
      'Do you know who killed him?'[3]   (Provotelle 1911: p. 88)

In Nafusi, this marker alternates with *ši*, indicating an Arabic borrowing:

(36) Fassato Nafusi: *aġr-ék-ši*   *agmâr?*
      (W Libya)   at-2M.SG-Q   horse
      'Do you have a horse?'   (Beguinot 1942: p. 139)

(37) *ad-i-nġû-n-ši?*
      IRR-1SG.ACC-kill-3M.PL-Q
      'Will they kill me?'   (Beguinot 1942: p. 113)

---

3. This example is translated as a positive question, but since Sened allows negation with *-š* alone, its interpretation is potentially ambiguous.

In Zuara and Djerba, however, it is *ša*, suggesting Berber origin; it is written as an independent word in available sources, but always seems to appear immediately following the verb:

(38) Zuara: *tə-ssə́n-əd*   *ša ləmmi mášəy ṣaləḥ a?*
      (W Libya) 2-know.PFV-2SG Q when   going Salih Q?
      'Do you know when Salih is going?'   (Mitchell 2009: p. 104)

(39) *yə-lla*   *ša sməʕan g-tiddárt a?*
      3M.SG-be.PFV Q Semʕan in-house Q?
      'Is Semʕan in the house?'   (Mitchell 2009: p. 97)

(40) Djerba: *rʼer-ouen cha midden eggeth g elḥʼoumeth   enn-ouen?*
      (S Tunisia) at-2M.PL Q   people many in neighbourhood GEN-2M.PL
      'Do you have many people in your neighbourhood?'
                                                    (Calassanti-Motylinski 1897: p. 382)

The distribution of this development strongly suggests that it started in Arabic and proceeded to Berber. In some regions it was calqued from Arabic, in others it was borrowed directly.

The Arabic clause-initial particle *waš* has likewise been borrowed directly into Saharan varieties near the Algerian-Moroccan border such as Figuig and Igli, as observed by Kossmann (2013: p. 305). At first sight, the clause-initial polar interrogative marker *ka* of Senhadja and Ghomara (northern Morocco) might likewise appear to derive from *kra*, as Kossmann suggests, in which case they could have been argued to be partial calques. However, in these languages *ka* also means 'if', from Arabic *kan*; interrogative and conditional *ka* are both non-spirant, while indefinite *ḵra* has a spirant *k* (Mourigh 2015: pp. 269, 347; Evgeniya Gutova p.c.). In neighbouring Tarifit, the clause-initial interrogative is *ma*, whose other senses include 'if' (Serhoual 2002: pp. 283–285). It thus appears more plausible to interpret *ka* as derived from 'if' than from *kra*.

6.   Calquing in Korandje

The close-knit Songhay language family developed in the Sahel and is almost entirely spoken there. While Berber and Arabic have had some impact on all Songhay varieties, their influence is profound only in Northern Songhay, and remains relatively minimal in southerly varieties such as Zarma. One variety, however, is spoken far enough north to be located within the Maghreb proper, deeply under the influence of northern Berber and Arabic: Korandje (Souag 2010). For perhaps 800 years, the speakers of Korandje have lived at the small oasis of Tabelbala in southwestern

Algeria; for 500 years or more, they were politically subordinate to neighbouring nomadic groups, notably the Arabic-speaking Ghenanma and the Berber-speaking Ait Atta, and their leading families claim Arab or Berber descent (Champault 1969; Souag 2015b). The extensive resulting influence on their language includes partial calquing of the grammaticalisation patterns of 'thing'.

6.1   The pre-contact situation

Throughout non-Northern Songhay, reflexes of *hàyà are used in the sense of 'thing':

(41)   Gao: *Sorko kul nga bis-ey   nda haya hinza*
                 Sorko all 3SG pass-3PL with thing three
       'All the Sorko surpass them by three things.'   (Prost 1956: p. 402)

(41)   Zarma: *hày-a   wò*
                 thing-DEF DEM
       'this thing'   (Sibomana 2008: p. 163)

In Korandje this usage is absent, but traces of it survive in other Northern varieties, cf. Tasawaq *hòò-ɣó* 'this thing' (Kossmann 2015: p. 92) and Tadaksahak *h-o* 'this (thing)' (Christiansen-Bolli 2011: p. 144). It must therefore be reconstructed for proto-Songhay.

All Songhay varieties, without exception, continue to use *hàyà* in the sense of 'something':

(42)   Gao: *haya g'   a   ra*
                 thing EXS 3SG LOC
       'There's something in it.'   (Prost 1956: p. 402)

(43)   Zarma: *kande ay   se   hay   hann-o*
                 bring 1SG DAT thing good-ADJ
       'Bring me something good.'   (Bernard & White-Kaba 1994: p. 151)

(44)   Korandje: *nə-ddzʊm haya nə-s-ddzʊm   haya?*
                 2SG-sow   thing 2SG-NEG-sow thing?
       'Did you sow anything, or did you sow nothing?'   (Souag 2010: p. 442)

yielding, under negation 'nothing':

(45)   Gao: *haya š'   a   ra*
                 thing NEG.EXS 3SG LOC
       'There's nothing in it.'   (Prost 1956: p. 402)

(46) Zarma: *a   si-nda   hay   fo*
      3SG NEG.EXS-with thing one
      'He has nothing.'   (Bernard & White-Kaba 1994: p. 151)

(47) Korandje: *a-s-ks=i.s   haya*
      3SG-NEG-leave=3PL.DAT anything
      'He left them nothing.'   (author's notes).

In many varieties, a slightly more lexical sense 'wealth, property' is also attested; it is an open question whether this is original, or derives from 'thing'. In either case, Proto-Songhay too seems to have covered the same core senses for **hàyà* as Arabic for *šayʔ* and Berber for **kắra*.

6.2   From indefinite pronoun to indefinite quantifier

Within Songhay, only one language has extended **hàyà* to quantifier use: Korandje, as briefly discussed in Souag (2010: p. 231). As a quantifier or as a pronoun, the reflex takes the form *hạ* in subject position and *haya* otherwise.

(48) Korandje: *lwərt   hạ   s-bạ*
      inheritance any NEG-EXS
      'There was no inheritance.'   (author's notes)

(49) *ndza man hạ   bạ…*
      if   fat any EXS…
      'If there's any fat …'   (Souag 2010: p. 232)

Comparable Arabic and Berber examples have already been seen above; cp:

(50) Ait Atta: *is t-uf-it   ka   n   usafar?*
      (SE Morocco) Q 2SG-find.PFV-2SG some GEN medicine.ANN?
      'Did you find any medicine?'   (Amaniss 1980: p. 746)

The influence of Arabic and Berber on Korandje is profound, and the distribution of this phenomenon within Songhay makes it impossible to interpret this as anything other than an example of that influence. However, what has been calqued is the colexification (François 2008) of 'any' with 'anything', rather than the whole construction. In Korandje *haya* follows what it quantifies, rather than preceding it as in Arabic and Berber. This corresponds to wider syntactic patterns: in Korandje, the specific indefinite article (homophonous with 'one'), and lower numerals in general, follow the noun, whereas in Arabic and Berber they precede it.

6.3   From indefinite pronoun to indefinite adverb

In Korandje, *hàyà* can be used as an indefinite adverb in negative and interrogative positions:

(51)   Korandje: *a-sə-bya*   *haya*
            3sg-neg-big   any
      'It's not big at all.'   (Souag 2010: p. 232)

This usage is infrequent in Korandje, and unattested anywhere else in Songhay. However, it does seem to be attested for Andalusi Arabic, and corresponds perfectly to what Jespersen's cycle would lead us to predict for the earlier stages of development in both Arabic and Berber.

Korandje has also borrowed the polar interrogative particle *waš* from western Maghrebi Arabic, but in this case there is no question of calquing.

7.   Conclusion

Prior to mutual contact, Arabic, Berber, and Songhay all shared the colexification of the noun "thing" and the indefinite pronoun 'something, anything, nothing'; only Arabic, however, additionally colexified this with an adverbial 'at all'. From the indefinite pronoun, many westerly Arabic dialects developed an indefinite quantifier 'some, any'; from the adverbial, it developed an emphatic negative marker and a polar question marker, both of which tended to become postverbal clitics (ultimately leading to the much-discussed postverbal negative marker). Under longstanding Arabic influence, some Berber varieties copied all of these developments. Korandje, under heavy influence from first Berber and later Arabic as well, copied the first two, both of which are attested in nearby Atlas Tamazight. Korandje, moreover, preserves a usage obsolete in modern North African Arabic, but already inferred from our knowledge of Jespersen's cycle: the use of 'thing' as an emphatic negative marker.

List of abbreviations

ACC	accusative	EXS	existential	NOM	nominative
ADJ	adjective	F	feminine	PASS	passive
ANN	annexed state	GEN	genitive	PFV	perfective
CENTRIF	centrifugal	INDF	indefinite	PL	plural
COP	copula	IPFV	imperfective	PTCP	participle
DAT	dative	IRR	irrealis	Q	question marker
DEF	definite	LOC	locative	SG	singular
DEM	demonstrative	M	masculine		
EP	epenthetic	NEG	negation		

References

Al-Jallad, A. (2014). What's a caron between friends? A Review Article of Wilmsen (2014), with Special Focus on the Etymology of Modern Arabic *ši*. *Bibliotheca Orientalis* 72: 1–2, 34–46.

Amaniss, A. (1980). Dictionnaire tamazight-français (parlers du Maroc centrale) [Tamazight-French Dictionary (dialects of Central Morocco)]. http://www.amazighworld.org/uploadedfiles/dictionnaire_Francais_Tamazight.pdf.

Basset, A. (1952). *La langue berbère* [The Berber Language]. (Handbook of African Languages pt. 1). London: Oxford University Press, for the International African Institute.

Beguinot, F. (1942). *Il berbero nefûsi di Fassâṭo: grammatica, testi raccolti dalla viva voce, vocabolarietti* [The Nefusi Berber of Fassato: Grammar, texts gathered orally, and lexica]. 2nd ed. Roma: Instituto per l'Oriente.

Behnstedt, P. (1994). *Der arabische Dialekt von Soukhne (Syrien)* [The Arabic dialect of Soukhne (Syria)]. Wiesbaden: Harrassowitz.

Behnstedt, P. (1996). *Die nordjemenitischen Dialekte. Teil 2: Glossar. Dal–Ghayn* [The North Yemeni dialects. Volume 2: Glossary: Dal-Ghayn]. (Jemen-Studien). Wiesbaden: Ludwig Reichert.

Ben Mamou, L. (2005). Dialecte Berbère de Tamezret (Tunisie) [Berber dialect of Tamezret (Tunisia)]. Atmazret. http://atmazret.info/ (3 November, 2006).

Benkato, A. (2014). The Arabic Dialect of Benghazi, Libya: Historical and Comparative Notes. *Zeitschrift für Arabische Linguistik* 59: 57–102.

Bernard, Y. and White-Kaba, M. (1994). *Dictionnaire Zarma-Français (République du Niger)* [Zarma-French Dictionary (Republic of Niger)]. Paris: Agence de coopération culturelle et technique.

Borg, A. (2004). *A Comparative Glossary of Cypriot Maronite Arabic (Arabic-English): With an Introductory Essay*. (Handbuch der Orientalistik 70). Leiden: Brill.

Brockett, A. A. (1985). *The spoken Arabic of Khābūra on the Bāṭina of Oman*. (Journal of Semitic Studies Monograph). Manchester: Journal of Semitic Studies, University of Manchester.

Brosselard, C. (1844). *Dictionnaire français-berbère (dialecte écrit et parlé par les Kabaïles de la division d'Alger)* [French-Berber dictionary (dialect written and spoken by the Kabyles of the Algiers division)]. Paris: Imprimerie Royale.

Brugnatelli, V. (2014). Berber negation in diachrony. In M.-B. Mosegaard Hansen & J. Visconti (eds.), 167–184. Amsterdam: John Benjamins.

Calassanti-Motylinski, G. A. de (1897). Dialogues et textes en berbère de Djerba [Dialogue and texts in Djerba Berber]. *Journal Asiatique*, 377–401.

Calassanti-Motylinski, G. A. de. (1905). Le Manuscrit arabo-berbère de Zouagha découvert par M. Rebillet : Notice sommaire et extraits [The Arabo-Berber manuscript of Zouagha discovered by M. Rebillet : Summary notice and extracts]. *Actes du XVIe Congrès des Orientalistes (Alger 1905)*, vol. II: 69–78. Paris.

Caubet, D. (1983). Quantification, négation, interrogation: les emplois de la particule "ši" en arabe marocain [Quantification, negation, interrogation: the uses of the particle "ši" in Moroccan Arabic]. *Arabica* 30(3): 227–245. https://doi.org/10.1163/157005883X00526

Caubet, D. (1993). *L'Arabe marocain* [Moroccan Arabic]. (Études Chamito-Sémitiques 1–2). Paris-Louvain: Éditions Peeters.

Champault, F. D. (1969). *Une oasis du Sahara nord-occidental : Tabelbala* [An oasis of the northwestern Sahara: Tabelbala]. Paris: CNRS.

Christiansen-Bolli, R. (2011). *A Grammar of Tadaksahak: A Berberised Songhay Language (Mali)*. (Berber Studies 31). Köln: Rüdiger Köppe.

Cotelle, H. (1847). *Le langage arabe ordinaire, ou, Dialogues arabes élémentaires, destinés aux français qui habitent l'Afrique* [Ordinary Arabic Language, or, Elementary Arabic dialogues, intended for the Frenchmen who live in Africa]. Alger: Dubos Frères.

Cowell, M. W. (1964). *A Reference Grammar of Syrian Arabic*. Washington: Georgetown University Press.

Davey, R. J. (2013). *Coastal Dhofārī Arabic: a sketch grammar*. Manchester: University of Manchester PhD thesis. https://www.escholar.manchester.ac.uk/uk-ac-man-scw:199131.

Delaporte, J. H. (1845). *Principes de l'idiome arabe en usage à Alger ; suivis d'un conte arabe avec la prononciation et le mot à mot interlinéaires* [Principles of the Arabic idiom used at Algiers; followed by an Arabic tale with pronunciation and interlinear glossing]. Paris: Hingray.

Destaing, E. (1907). *Etude sur le dialecte berbère des Beni-Snous* [Study of the Berber dialect of the Beni-Snous]. Paris: Ernest Leroux.

Diem, W. (2014). *Negation in Arabic: A Study in Linguistic History*. Wiesbaden: Harrassowitz.

Feghali, M. (1919). *Le parler de Kfar 'Abida (Liban-Syrie). Essai linguistique sur la phonétique et la morphologie d'un parler arabe moderne* [The variety of Kfar 'Abida (Lebanon-Syria). Linguistic study of the phonetics and morphology of a modern Arabic dialect]. Paris: Ernest Leroux.

François, A. (2008). Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In M. Vanhove (ed.), *From Polysemy to Semantic Change: Towards a typology of lexical semantic associations*. Studies in Language Companion Series, vol. 106 (pp. 163–215). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.106.09fra

Grigore, G. (2009). *L'arabe parlé à Mardin: monographie d'un parler arabe "périphérique"* [THe Arabic spoken at Mardin: Monograph on a "peripheral" Arabic dialect]. Piscataway, NJ: Gorgias Press.

Haspelmath, M. and Caruana, J. (1996). Indefinite pronouns in Maltese. *Rivista di Linguistica* 8(1): 213–227.

Heine, B. & Kuteva, T. (2005). *Language Contact and Grammatical Change*. (Cambridge Approaches to Language Contact). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511614132

Jespersen, O. (1917). *Negation in English and Other Languages*. København: A. F. Høst.

Khamed Attayoub, A. (2001). La tetserrét des Ayttawari Seslem: Identification socio-linguistique d'un parler berbère non documenté parmi les Touaregs de l'Azawagh (Niger) [The Tetserrét of the Ayttawari Seslem: Sociolinguistic identification of an undocumented Berber variety spoken among the Tuaregs of the Azawagh (Niger)]. INALCO, Paris MA thesis.

Kossmann, M. (1999). *Essai sur la phonologie du proto-berbère* [Study of Proto-Berber phonology]. Köln: Rüdiger Köppe.

Kossmann, M. (2013). *The Arabic Influence on Northern Berber*. Leiden: Brill. https://doi.org/10.1163/9789004253094

Kossmann, M. (2015). A Tasawaq (Northern Songhay, Niger) Text with Grammatical Notes. *Linguistic Discovery* 13(1): 82–118. https://doi.org/10.1349/PS1.1537-0852.A.448

Lahdo, A. (2009). *The Arabic dialect of Tillo in the region of Siirt (South-eastern Turkey)*. (Acta Universitatis Upsaliensis. Studia Semitica Upsaliensia). Uppsala: Uppsala Universitet.

von Landberg, C. (1901). *Etudes sur les dialectes de l'Arabie Méridionale*. Vol. I [Studies on the dialects of Southern Arabia. Vol. I]. Leiden: Brill.

Lanfry, J. (1968). *Ghadamès I: Textes : Notes philologiques et ethnographiques* [Ghadames I. Texts: Philological and ethnographical notes]. Fort-National (Algérie): Le Fichier Périodique.

Lanfry, J. (1973). *Ghadamès II: Glossaire (parler des Ayt Waziten)* [Ghadames II: Glossary (Ayt Waziten variety)]. Fort-National (Algérie): Le Fichier Périodique.

Lucas, C. (2010). The development of negation in Arabic and Afro-Asiatic. Cambridge: Emmanuel College. University of Cambridge. PhD thesis.

Lucas, C. (2013). The development of negation in Afro-Asiatic languages. In Anne Breitbarth, C. Lucas & D. Willis (eds.), *The History of Negation in the Languages of Europe and the Mediterranean*, vol. 1: 399–452. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199602537.003.0010

Lucas, C. and Lash, E. (2010). Contact as catalyst: The case for Coptic influence in the development of Arabic negation. *Journal of Linguistics* 46(2): 379–413. https://doi.org/10.1017/S0022226709990235

Lux, C. (2013). *Le tetserrét, langue berbère du Niger : Description phonétique, phonologique et morphologique, dans une perspective comparative* [Tetserrét, a Berber language of Niger: Phonetic, phonological and morphological description in a comparative perspective]. (Berber Studies 38). Köln: Rüdiger Köppe.

Madouni-La Peyre, J. (2003). *Dictionnaire Arabe Algérien-Français: Algérie de l'ouest* [Algerian Arabic-French Dictionary: Western Algeria]. (Dictionnaires L & M). Paris: Langues & Mondes – L'Asiathèque.

Mitchell, T. F. (2009). *Zuaran Berber (Libya) : Grammar and Texts*. (ed.) H. Stroomer & S. Oomen. (Berber Studies v. 26). Köln: Rüdiger Köppe.

Mourigh, K. (2015). *A Grammar of Ghomara Berber*. Leiden: University of Leiden.

Nadwi, A. A. (1968). A study of the Arabic dialects of the Belad Ghamid and Zahran region of Saudi Arabia on the basis of original field recording and an examination of the relationship to the neighbouring regions. Leeds: University of Leeds PhD thesis. http://etheses.whiterose.ac.uk/918/ (20 June, 2016).

Nicolas, F. (1953). *La langue berbère de Mauritanie* [The Berber language of Mauritania]. (Mémoires de l'Institut Français d'Afrique Noire 33). Dakar: Institut Français de l'Afrique Noire.

Obler, L. K. (1990). Reflexes of Classical Arabic *šay'un* in the dialects: Synthetic forms of language change. In James Bellamy (ed.), *Studies in Near Eastern Culture and History in Memory of Ernest T. Abdel-Massih*. Ann Arbor, MI: University of Michigan Press.

Owens, J. (1984). *A Short Reference Grammar of Eastern Libyan Arabic*. Wiesbaden: Otto Harrassowitz.

Prost, A. (1956). *La langue soṅay et ses dialectes* [The Soṅay language and its dialects]. Dakar: IFAN.

Provotelle, P. (1911). *Etude sur la tamazir't ou zénatia de Qalâat Es-Sened* [Study of the Tamazir't or Zenatia of Qalaat Es-Sened]. Paris: Ernest Leroux.

van Putten, M. (2014). *A Grammar of Awjila (Libya) Based on Paradisi's Work*. Köln: Rüdiger Köppe.

Serhoual, M. (2002). *Dictionnaire Tarifit-Français* [Tarifit-French Dictionary]. Tetouan: Université Abdelmalek Essaâdi PhD thesis.

Sibomana, L. (2008). *Le Zarma parlé : esquisse grammaticale, lexique, textes* [Spoken Zarma: Grammar sketch, lexicon, texts]. Berlin: Lit.

Singer, H.-R. (1958). Neuarabische Fragewörter. Ein Beitrag zur historischen und vergleichenden Grammatik der arabischen Dialekte [Neo-Arabic question words. A contribution towards the historical and comparative grammar of the Arabic dialects]. Erlangen: Friedrich-Alexander-Universiteit dissertation.

Souag, L. (2009). Siwa and its significance for Arabic dialectology. *Zeitschrift für Arabische Linguistik* 51: 51–75.

Souag, L. (2010). Grammatical Contact in the Sahara: Arabic, Berber, and Songhay in Tabelbala and Siwa. School of Oriental and African Studies, University of London. PhD thesis.

Souag, L. (2015a). Archaic and innovative Islamic prayer names around the Sahara. *Bulletin of the School of Oriental and African Studies* 78(2): 357–374.
https://doi.org/10.1017/S0041977X15000075

Souag, L. (2015b). Explaining Korandjé: Language contact, plantations, and the trans-Saharan trade. *Journal of Pidgin and Creole Studies* 30(2): 189–224. https://doi.org/10.1075/jpcl.30.2.01sou

Souag, L. & van Putten, M. (2015). Attrition and revival in Awjila Berber. *Corpus* 14: 23–58.

Stumme, H. (1896). *Grammatik des tunisischen Arabisch, nebst Glossar*. Leipzig: Henrichs.

Taine-Cheikh, C. (2008). *Dictionnaire zenaga-francais : le berbere de Mauritanie présenté par racines dans une perspective comparative* [Zenaga-French dictionary: The Berber of Mauritania in a comparative perspective]. (Berber Studies v. 20). Köln: Rüdiger Köppe.

Torchon (1871). Essai d'une grammaire chaouie [Attempt at a Chaoui grammar]. In A. Sierakowski (ed.), *Das Schaui, ein Beitrag zur Berberischen Sprachen und Völkerkunde*. Dresden: Kraszewski.

Wilmsen, D. (2014). *Arabic Indefinites, Interrogatives, and Negators: A Linguistic History of Western Dialects*. (Oxford Studies in Diachronic and Historical Linguistics 14). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198718123.001.0001

Woodhead, D. R. & Beene, W. (1967). *A dictionary of Iraqi Arabic, Arabic-English*. Washington, DC: Georgetown University Press.

Yoseph, J. (2012). *Der arabische Dialekt von Mharde (Zentralsyrien)* [The Arabic dialect of Mharde (Central Syria)]. (Semitica Viva 51). Wiesbaden: Harrassowitz.

Arabic and Berber in contact

Arabic in a minority situation in Al Hoceima region

Dominique Caubet

LaCNAD, Centre Jacques Berque

Near the Berber-speaking town of Al Hoceima, there are a few hamlets and villages where people speak Arabic and find themselves in a situation where Berber is the dominant language. These dialects of Moroccan Arabic have seldom been described. What is taking place on the border between Berber and Arabic in this region? What types of contact? What influences? We visited one village on the Berber speaking side (Taounil) and one hamlet on the Arabic-speaking side. Our fieldwork was tentative transdisciplinary work by linguists and ethnobotanists, which allowed us to collect very spontaneous data, since the stress was put on the ethnobotanic questioning. We present here our results, analysing the specific traits of these dialects.

1. Introduction: A border region

Near Berber speaking Al Hoceima, there are a few hamlets and villages (*duwwar* or *dechar*), where people speak Arabic and find themselves in a minority situation. These dialects of Moroccan Arabic have seldom been described.

The data used here is part of a larger project concerning the Arabic dialects spoken in the North West of Morocco, mostly in what is called the Jbala region. The Jbala dialects are fairly different from the dialects spoken in Central Morocco, they present a series of specific traits that were described by dialectologists in the beginning of the 20th century and classified as belonging to the first layers of Arabisation of Morocco.[1] Since these dialects sounded different and were sometimes mocked outside the region, it was predicted that they would disappear and be absorbed in a more central koine. But we witnessed that they are still being spoken nowadays.

---

1. See Colin (1921), Lévy-Provençal (1922), Marçais (1911) etc.

Unpublished studies directed by Simon Lévy in the early 90's[2] (that we accessed in 2012) showed that there had been little evolution in over a period of over sixty years full of major events that could have impacted massively people's speech, such as colonisation by Spain and France, imposition of new languages (Spanish and French), nationalism, decolonisation, Arabisation etc. We decided to revisit some of these places, and among them, a border region only described in the unpublished work of Maghdad for her Mémoire de licence (see Maghdad 1993 and Caubet 2017). Simon Lévy (1998: p. 12) defined it as:

> Farther east, not far from Alhucemas, the tribe of Beni Yitteft is Riffian, embedded between Bokoya and Ait Ouriaghel – who speak a Riffian dialect (*ed. Berber*), is half Arabized. Their dialect was recently studied by one of our students (*ed. Maghdad 1993*). It is a dialect with Jebli features, strongly influenced by the Riffian spirant substrate (*ed. Berber*).[3]

This is a border region where two languages (Berber and Moroccan Arabic) have been tangled since the 8th century and have evolved in very close contact, borrowing from each other on all levels of language, phonetic, morphosyntactic and lexical. Nowadays bilingualism is very common, mostly on the part of Berberophones.

For a detailed discussion on the language in 2012–2015, see Section 6.

Amédée Renisio, in his 1932 study of *Dialectes berbères des Beni Iznassen, du Rif et des Senhadja du Srair*, published a map of the tribes and drew a line of the limit between Berber and Moroccan Arabic. It does not seem to have changed for our area, some eighty years later. We did our fieldwork exactly across the line between Beni Iṭṭeft (both Arabic and Berber-speaking, as S. Lévy was mentioning) and Boqqoya (Berber-speaking) tribes.

---

2.   I discovered these studies in a footnote of article by Simon Lévy (Lévy 1998: p. 12 note 6). They were "Mémoires de licence" in the Spanish Department of Mohamed V University where Simon was a Professor.

3.   My translation from French: "Plus à l'est, non loin d'Alhucemas, la tribu de Beni Yitteft, rifaine, enfoncée entre Bokoya et Ait Ouriaghel, au parler tarifit, est à moitié arabisée. Son parler a été récemment étudié par une de nos étudiantes. C'est un parler aux traits jebli, fortement personnalisé par le substrat rifain spirant".
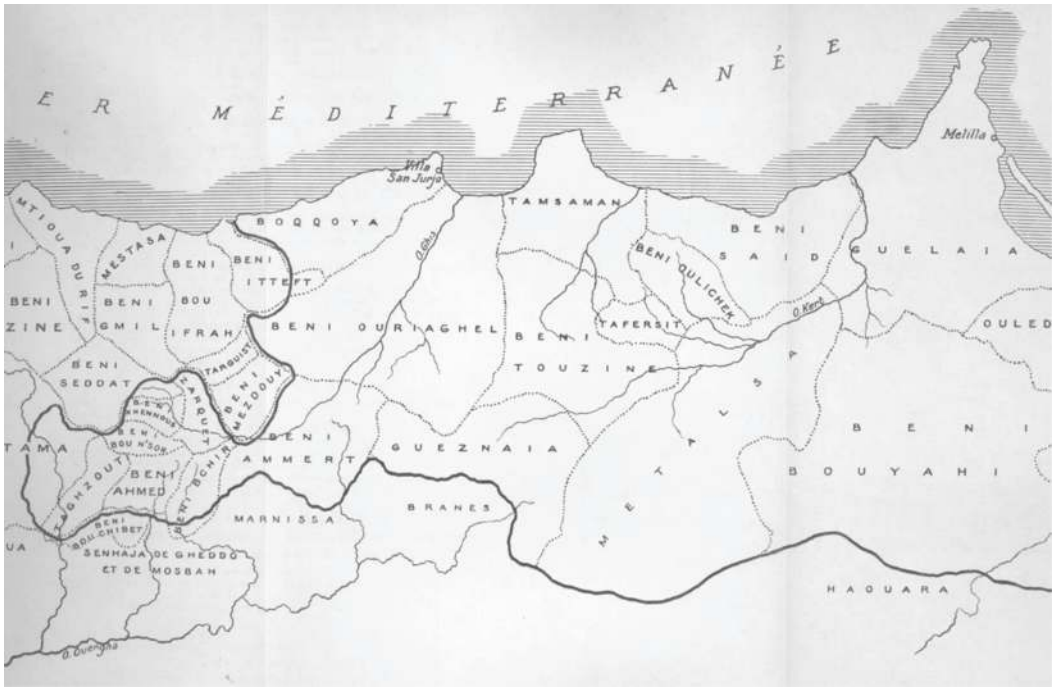
Map 1. Renisio's 1932 *Dialectes berbères des Beni Iznassen, du Rif et des Senhadja du Srair* – Villa San Jurjo is the former name of Al Hoceima (Renisio 1932)

2.   Beni Iṭṭeft "Revisited"[4]…

We "revisited" the region in February 2014 with the PICS programme *La Montagne et ses savoirs*,[5] and Centre Jacques Berques's 'Programme Jbala', with ethnobotanist Yildiz Thomas,[6] S. Lévy's former student, Amal Maghdad and a Master's student from Oujda University, Khalid El Jattari. We stayed in Al Hoceima National Park in the village of Taounil in the heart of the Boqqoya tribe.

The Beni Iṭṭeft (Ait Iṭṭeft in Berber) consider themselves as Ryafa (Riffians), not Jbala, although the large majority speak Arabic, as shown on Renisio's map, where only a small fraction to the South-East speaks Berber (Map 2). The language borders do not seem to have changed since then.

---

4.   This article aims at "revisiting" the first unpublished description done by A. Maghdad in 1993 under S. Lévy's supervision.

5.   PICS CNRS-CNRST 2013–2015 – IREMAM – Université de Fès, Tétouan.

6.   CNRS- UMR CEFE 5175 Montpellier.

Map 2. A detail of Renisio's map for *Dialectes berbères des Beni Iznassen, du Rif et des Senhadja du Srair* – Villa San Jurjo is the former name of Al Hoceima (Renisio 1932)

In 1968, Maurer (1968: p. 15) drew a map of the Riffian tribes -among whom the Beni Iṭṭeft (and the Arabic speaking Mtioua, Mestasa and Bni Boufrah) – and of their border with the Sanhaja-de-Sraïr and further to the West, the Ghomara (see Map 3).

According to Maurer (Map 4, 02–1–01), the Beni Iṭṭeft tribe counts four fractions, three of which spoke Arabic in the 1960's, El-Amaïr, Izeroualène and El Ouadiyne (Maurer's spelling), and the Aït Aïssa fraction who mostly spoke Berber (also see Map 5 from Renisio 1932).

The place we re-visited, *Msek,* is the only Arabic speaking *duwwar (deššar)* in the Aït Aïssa fraction. In her dissertation, Amal Maghdad (1993: p. 6) defined Msek as "un pequeño y único núcleo arabiziado dentro de una fracción (Aït Aïssa) del Rif central" (a little and unique arabized nucleus inside a fraction (Aït Aïssa) of the Central Rif).

Msek thus appears to be *a minority* (Arabic speaking among the Berber speaking Aït Aïssa fraction), *inside another minority*, the Aït Aïssa fraction, as opposed to the mostly Arabic speaking Beni Iṭṭeft tribe.

Map 3. Extract from a map by Maurer (Fig. 3 Les tribus dans les montagnes du Rif central) (Maurer 1968: p. 15)

The Arabic dialect of Msek is a Prehilali dialect which shares a number of traits with the Jbala, but it also has its own characteristics. It has been in close contact with a *Rifi* dialect of Berber, i.e. a Zeneta variety and not a Senhaja one,[7] for centuries.

3. Method

In Msek we re-visited one of the families Maghdad (1993) had recorded 22 years previously for her initial 1992 fieldwork. We worked there in February 2014 with a mother of about 40, S., and her son Y., who was 13 at the time.

---

7. See Kossman (2017a).

Map 4. Extract from Maurer (1968: 16) (Fig. 4 Communes rurales et fractions dans le Rif central)
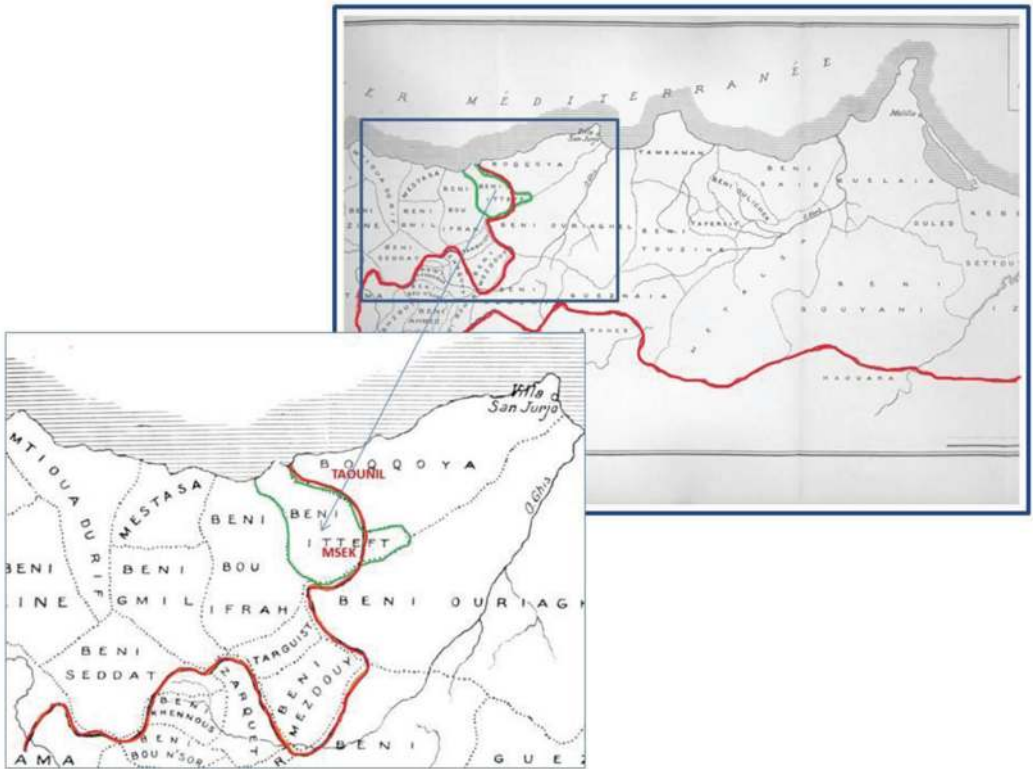
Map 5. Extract from Renisio (1932). The limits of Berberophony (red line taken from the limit proposed by Renisio). The Bni Itteft (green line) with a Berber-speaking part which corresponds to the tribe of Aït Aïssa of which Msek is part, located exactly to the West of this border. Villa Sanjurjo is the current location of the town of El Hoceima

In Taounil, our informant was our host, F., a woman of 50 from the Boqqoya tribe; she has never been to school, her native language is Berber and she told us she learned Moroccan Arabic, as a second language, from her neighbours from the Arabic speaking fractions of the Beni Iṭṭeft – probably Izeroualène or El-Amaïr (see Map 4) and from the Beni Boufrah. She also spent 15 years working with a Moroccan Riffian family living in Ceuta; she seldom went out, but she may have picked up some expressions from Ceuta.

When asked if the Beni Iṭṭeft learn Berber when living among the Boqqoya, she says humorously: *huma kayhəḍṛu l-ɛarbiya waxxa huma rifiyin, huma kayhəḍṛu l-ɛarbiya* "they speak Arabic, even if they are Riffians, they speak Arabic", which shows a dominant attitude on their part, different from what happens in Msek.

3.1   A rare situation: Arabic as a minority language

In Msek, when I asked the boy – in the presence of his mother – what languages he spoke, he answered first: *l-ɛarbiya* (Arabic); when I asked whether he spoke Berber, he said *šwiya* (a little) and his mother promptly interrupted saying: *la, ts, ḥna ma kanhəḍṛu š š-šəlḥa!* "No! Ts! We don't speak Berber here!".

She was stating clearly their linguistic identity in a minority situation. When we rephrased the question, asking with whom the boy spoke Berber, he answered: *f-əṭ-ṭriq dyal l-mədrasa, f-əṭ-ṭobis!* "On the way to school, on the bus". He had to learn Berber because he was going to school in a village with the Aït Aïssa and Berber was the language spoken in that environment (when playing soccer, during the intervals, on the way to school, all the exchanges took place in Berber…). The boy finds himself in a situation which is sufficiently unusual to be noticed, where Arabic is a minority language and Berber the dominant one.

We will try and compare these two mirrored situation: when Moroccan Arabic (M.A.) is a minority language in Msek, and when it is a second language which an illiterate woman learned from her neighbours in order to socialize in her own village of Taounil.

3.2   Transdisciplinarity

The fieldwork in both places was tentative transdisciplinary work between linguists and ethnobotanists; it proved very productive although we had to adapt to our respective enquiry methods. The linguistic material we collected was completely natural and fluent because the informants were answering questions about basic techniques of picking or preparing. We discussed *zembu* (barley paste)*, tasukkant*

(wild asparagus) and *beqqula* (varieties of spinach) in Taounil; and various types of oil, almonds, cereals and bread in Msek.

Short texts will be presented complete and the traits resulting from contact will be discussed and compared in 6: *tasukkant* and *zembu* for Taounil, *ɛalwana* "baked olive oil" and bitter almonds oil for Sefri (Msek). The questioning on these practices was initiated by ethnobotanist Yildiz Thomas.[8]

4. Taounil data: *Tasukkant* (wild asparagus) and *zembu* (young barley paste)

When we stayed in Taounil, we were lucky to be present – at the end of February 2014 – for the very short season of wild asparagus, *tasukkant*.

4.1   Tasukkant

4.1.1   *The word 'tasukkant', a loanword?*
As a preliminary remark, we will comment on the name itself. The word *ta-sukkan-t* is a feminine substantive in Berber. Wild asparagus are called *səkkum (əl-bərr)* in M.A. It can be found in Mercier's and Colin's dictionaries:[9] Why a feminine word in Berber? It is curious to note that the word has feminine agreement in text 1, and in the masculine text 2 (see below, 6.6 for agreement and 6.8 for the lexicon).

Mercier gives: "SEKKUM, asperge; – el-berr, asperge sauvage" (wild asparagus), *berr* meaning 'sauvage (legume, fruit)' (wild for a vegetable or a fruit); Colin: "*səkkūm, n. coll. Bot.* Asperges sauvages" (wild asparagus).

It is not clear whether Berber borrowed from Arabic or vice-versa. In a discussion with Salem Chaker, he said the root existed both in Berber – but not in the Rif – and in Arabic. He explained the passage from *sekkum* to *ta-sukkan-t* as follows: "*m* becomes *n* before the *-t* suffix, by assimilation to the following dental".[10]

We will see a similar phenomenon with the term *tažəfnit* for Msek (text 4 in 5.1 and 6.8). All the forms that will be analysed in Section 6 are in bold in the texts.

4.1.2   *The texts*
We will give successively the text and its translation, and the detailed gloss.

---

8. See Caubet & Thomas and Thomas & Caubet (2017).

9. See Mercier (1951: p. 182) and for Colin in Iraqui-Sinaceur (1993 vol. 4, p. 831).

10. Thanks to Salem Chaker (personal communication).

Text 1:   *ka-tə-nbət* **hayda** *f-əl-aṛḍ, hiya fi-ha š-šukk, bəllati, hiya ġir hnaya fuq-d-daṛ*
*dyal-na, hiya ka-t-kun ġiṛ f-had-əl-***waxt** *(weqt), ka-tənbət b-waḥd-***a.**
It grows like this, in the ground, it has thorns, wait, you can find it right here,
above our house, it grows only at this time of year, it grows by itself.

Text 1   *ka=tə-nbət*   ***hayda***   *f-əl=aṛḍ*
PRVB=3F.SG-grow\IPFV   ADV   PREP-DEF=N.F.
'It grows like this, in the ground'
hiya   *fi-ha*   *š=šukk*   *bəllati*
PRO.IDP.3F.SG   PREP-OBL.3F   DEF=N.M.   ADV
'It has thorns, wait'
hiya   *ġir*   *hnaya*   *fuq-d=daṛ*   *dyal=na*
PRO.IDP.3F.SG   CONJ   ADV.LOC   PREP-DEF=N.M.   POSS=OBL.1PL
'you can find it right here, above our house…'
hiya   *ka=t-kun*   *ġiṛ*   *f-had=əl=***waxt** *(weqt)*
PRO.IDP.3F.SG   PRVB=3F.SG-be\IPFV   CONJ   PREP-DEM.PROX=DEF=N.M
'it grows only at this time of year'
ka=tə-nbət   *b-waḥd=***a**
PRVB=3F.SG-grow\IPFV   PREP-NUM=OBL.3F.SG
'it grows by itself.'

Text 2:   *Tasukkant? f-əl-***maᷠ,** **ka-nsəlquw-ăh** *f-əl-***maᷠ** *ka-yṭeb, məlli ka-y-ṭeb…*
ka-nɛəṣṣṛuw-ăh ka-nɛəmluw-ăh *yəqtăṛ,* **dik-əs-saɛa ka-nṭəyybuw-ăh,**
*ka-nɛəml-u l-***u** *ət-tăwm, ka-nɛəml-u l-***u** *əl-qəsbuṛ, u* **ka-nɛəmluw-ăh** *f-əl-măqla,*
*dik-s-saɛa ka-nɛəml-u l-***u** *əl-băyṭat, dik-əs-saɛa, ka-yə… ka-yəntkəl.*
Asparagus? In water, we boil it, in water, it cooks, when it is cooked, we press
it and we put it, to drip, and then we prepare it; we add garlic to it, we add
coriander to it; and we put it in the frying pan, and then, we add eggs to it (the
preparation), and then, you can eat it.

Text 2   *tasukkant*   *f-əl=***maᷠ**   *ka=n-səlq-uw=ăh*
N.F.   PREP-DEF=N.M   PRVB=1-boil\IPFV-PL=OBJ.3M.SG
'Asparagus? In water, we boil it…'
*f-əl=***maᷠ**   *ka=y-ṭeb*   *məlli*   *ka=y-ṭeb…*
PREP-DEF=N.M   PRVB=3M.SG-cook\IPFV   CONJ   PRVB=3M.SG-cook\IPFV
'…in water, it cooks, when it is cooked…'
ka=n-ɛəṣṣṛ-uw=ăh   *ka=n-ɛəml-uw=ăh…*
PRVB=1-press\IPFV-PL=OBJ.3M.SG   PRVB=1-do\IPFV-PL=OBJ.3M.SG
'we press it and we put it…'
yəqtăṛ   *dik=əs=saɛa*   *ka=n-ṭəyyb-uw=ăh*
3M.SG-drip\IPFV   DEM.DIST=DEF=N.F.   PRVB=1-cook\IPFV-PL=OBJ.3M.SG
'…to drip, and then we prepare it;'

> *ka=n-ɛəml-u*　　　　*l-u*　　　　　　*ət=tăwm*
> PRVB=1-do\IPFV-PL　PREP-OBL.3M.SG　DEF=N.M.
> 'we add garlic to it'
> *ka=n-ɛəml-u*　　　　*l-u*　　　　　　*əl=qəsbuṛ*
> PRVB=1-do\IPFV-PL　PREP-OBL.3M.SG　DEF=N.M.
> 'we add coriander to it'
> *…u*　　**ka=n-ɛəml-uw=ăh**　　　　*f-əl=măqla*
> …and　PRVB=1-do\IPFV-PL=OBJ.3SG.M　PREP-DEF=N.F
> '… and we put it in the frying pan'
> *dik=s=saɛa*　　　　*ka=n-ɛəml-u*　　　*l-u*　　　　*əl=băyṭat*
> DEM.DIST=DEF=N.F.　PRVB=1-do\IPFV-PL　PREP-OBL.3SG.M　DET=N.F.PL
> 'and then, we add eggs to it (the preparation)'
> *dik=əs=saɛa*　　　*ka=yə…*　*ka=yə-ntkəl.*
> DEM.DIST=DEF=N.F.　PRVB=3　PRVB=3-eat\IPFV\PASS
> 'and then you …, you can eat it.'

Tasukkant, *asparagus acutifolius,* grows in specific territories and particularly on the southern slopes, near the Mediterranean, which corresponds exactly to the situation of Taounil.

4.2　Zembu

Zembu is the name of the preparation which can be made from several cereals, depending on what is available in the region. In Taounil, it is thin flour made from roasted young (green) barley. It is called *zembu* in Al Hoceima region, *tazemmit* near Nador, which is a loanword in Berber, coming from *zəmmēṭa,* the name given to the preparation in the Jbala area, near Ouazzane.

In Taounil, when you want to eat it, you make a paste with warm water, forming small bowls; you eat it with butter, oil or honey. Once the flour is grounded and sieved, you get thin flour (*zembu*) and larger grains (*dšiša*) that can be prepared like couscous. F. made it for us and commented:

Text 3:　*kanġərbəl zəmbu daba, hayda, šuf!* **ha huwa əd-dšiša,** *ha əd-dšiša!* **zəmbu ha hiya, ha huwa!** *hadi əd-dšiša u* **hadi zəmbu.**
　　zəmbu: haḏi ka-yaklu-ha **hayḏa,** *ka-yɛəžn-u-ha* **b-əl-ma sxunin** *u kayaklu-ha, (h)a-hi b-əz-zəbda, (h)a-hi b-əz-zīt, (h)a-hi b-lə-ɛsəl, lli bġiti, lli bġiti* **ntina.**
　　u **hadak ka-nṭayybuw-ăh** *ka-nɛəmlu bi-ha lə-ḥrīra wəlla* **ka-nfowwṛuw-ăh** *mɛa-l-lḇən wəlla* **ka-nṭayybuw-ăh hayda, ka-yəntkəl hayda** *ka-yɛəml-u l-u z-zīt u* **ka-yakluw-ăh** *wəlla* **ka-yɛəmlu-h** *bḥāl səksu ɛāwəd u ka-yɛəml-u ɛli-h* **l-lḥœm,** *wəlla mɛa* **ǧ-ǧdād,** *wəlla bḥal səksu.*

I'm sifting *zembu* now, like this, look! This is *dchicha,* here is *dchicha*! *Zembu*, here it is (fem), here it is (masc)! This one is (fem) *dchicha* and this one is (fem) *zembu*!

This one (fem. zembu), they eat it like this, they knead it with hot water, and they eat it, either with butter … or with oil, or with honey, whatever you want, whatever you prefer

and the other one (masc.), we cook it, we make *harira* (Ramadan soup) with it, or we steam it (and eat it) with buttermilk, or we cook it like this, you can eat it like this, we put it with oil and they eat it, or they prepare it like couscous too, and they put meat with it, or with chicken, or, like couscous.

Text 3 *ka=n-ġərbəl   zəmbu   daba   hayda   šuf*
PRVB=1-sift\IPFV N.M.   ADV   ADV   look\IMP.M
'I'm sifting zembu now, like this, look!'
ha   huwa   əd=dšiša   ha   əd=dšiša
PRST PRO.IDP.3M.SG DEF=N.F.   PRST DEF=N.F.
'This is dchicha, here is dchicha!'
zəmbu   ha   hiya   ha   huwa
N.M.   PRST PRO.IDP.3F.SG PRST PRO.IDP.3M.SG
'zembu, here it is (fem), here it is (masc)'
hadi   əd=dšiša   u   hadi   zəmbu
PROX.F   DEF=N.F.   CONJ PROX.F N.M.
'This one is (fem) dchicha and this one is (fem) zembu'
haḏi   ka=y-akl-u=ha   hayḏa
PROX.F   PRVB=3-eat\IPFV-PL=OBJ.3F.SG   ADV
'This one (zembu), they eat it like this'
ka=y-ɛəžn-u=ha   b-əl=ma   sxun-in
PRVB=3-knead\IPFV-PL=OBJ.3F.SG   PREP-DEF=N.M ADJ.PL
'they knead it with hot water'
u   ka=y-akl-u=ha   (h)a=hi   b-əz=zəbda
CONJ PRVB=3-eat\IPFV-PL=OBJ.3F.SG   PRST=3F.SG PREP-DEF=N.F.
'and they eat it, either with butter…'
(h)a=hi   b-əz=zīt   (h)a=hi   b-lə=ɛsəl
PRST=3F.SG PREP-DEF=N.F. PRST=3F.SG PREP-DEF=N.M.
'… or with oil, or with honey…'
lli   bġi-ti   lli   bġi-ti   ntina
REL want\PFV-2F.SG REL want\PFV-2F.SG PRO.IDP. 2SG
'Whatever you want, whatever you prefer'
u   hadak   ka=n-ṭayyb-uw=ăh
CONJ DIST.M.SG PRVB=1-cook\IPFV-PL=OBJ.3M.SG
'and the other one, we cook it'

ka=n-ɛəmlu   *bi-ha*   *lə=ḥrīra*
PRVB=1-do\IPFV-PL PREP-OBL.3F.SG DEF=N.F
'we make 'harira' (Ramadan soup) with it'

wəlla   ***ka=n-fowwṛ-uw=ăh***   *mɛa-l=lḅən*
CONJ   PRVB=1-steam\IPFV-PL=OBJ.3M.SG PREP-DEF=N.M.
'or we steam it (and eat it) with buttermilk'

wəlla   ***ka=n-ṭayyb-uw=ăh***   *hayda*
CONJ   PRVB=1-cook\IPFV-PL=OBJ.3M.SG ADV
'or we cook it like this'

ka=yə-ntkəl   *hayda* *ka=y-ɛəml-u*   *l-u*
PRVB=3-eat\IPFV.PASS ADV   PRVB=1-do\IPFV-PL PREP-OBL.3M.SG
'you can eat it like this, we put it with…'

z=zīt   *u*   *ka=y-akl-uw=ăh*
DET=N.M   CONJ   PRVB=3-eat\IPFV-PL=OBJ.3M.SG
'…oil and they eat it'

wəlla   ***ka=y-ɛəml-u=h***   *bḥal* *səksu*
CONJ   PRVB=3-do\IPFV-PL=OBJ.3M.SG ADV   N.M.
'or they prepare it like couscous…'

ɛāwəd *u*   *ka=y-ɛəml-u*   *ɛli-h*   ***l=lḥœm***
ADV   CONJ   PRVB=3-do\IPFV-PL PREP-OBL.3M.SG DEF=N.M
'… too, and they put meat with it'

wəlla   *mɛa*   ***ǧ=ǧdād***   *wəlla* *bḥal* *səksu*
CONJ   PREP   DEF=N.M   CONJ   CONJ   N.M.
'…or with chicken, or, like couscous.'

We will analyze the contact phenomena and the linguistic peculiarities in paragraph 6.

5.   Msek data

We discussed several techniques in Sefri (Msek), but Yildiz Thomas asked S. if (and how) she made a special type of oil, from baked olives (called *ɛalwana*), and with bitter almonds. They both involved the very particular use of a special cloth, *dərra d-əl-ḥayati*.

5.1   Ɛalwana

This is a special olive oil, made from olives previously left to dry all night in the oven, which gives it a very peculiar smoked taste.

Text 4: ka-n-žni-w *z-ze^{y}tun*, **ka-nžibu-h**, ka-nḥmiw **l-fəṛṛan**, ka-nḥmiw, ka-yttəḥma
məzyan, bḥal lli ġadi ṭṭiyyəb əl-xobz; **ka-nməsḥu-h** məzyan, n-žib-u z-zitun,
nnəqqiw-əh wə **nɛəddl-u** fi-h, **nɛəddl-u** fi-h y-bat l-lila kamla, yibəs, ṭḥəss bi-h
yabəs; f-əṣ-ṣbaḥ, məlli tḥăll əl-fəṛṛan, tṣib-u yabəs, thəzz-u tžib əl-məhraz, tdoqq-u,
tdoqq-u məzyan ḥta yetdəqq **dik-lə-ɛṭam** dyal-u məzyan, məlli **ndoqqu-h** nžib-u
hadikəə…**tažəfnit**, baš ka-nɛəžn-u, ḥna ka-**nqulu**-ha **tažəfnit.**

bqa tdoqq u təɛməl fi-ha, doqq u ɛməl fi-ha, žib əl-faxəṛ, **ɛəml**-u f-əl-məžmaṛ,
šḥəl (šɛəl) əl-ɛafiya u nəzzəl ɛli-ha **dik-tažəfnit**, **dik-əl-gəṣea**, xalli-h ḥta yəsxun
məzyan u bda ɛəžn-u b-yəddə-k, w-**əntina** ka-tšuf-u ka-yxəṛṛəž z-zit, ka-tɛəžn-u
hayda, hayda, ṣafi u tžib **dərra d-əl-ḥayati** u ɛəṣṣəṛ; ṣafi u **z-zit ka-txrəž** məzyan!
We collect the olives, we bring them, we heat the oven, we heat it, until it is
very hot, as if you were going to bake bread, we clean it off well, we bring the
olives, and we clean them, and we work on it, we prepare it and it spends the
whole night (in the oven) in order to make it dry, you make sure it is dry, the
next morning, when you open the oven and you find them dry, you take them
out, you bring the mortar and pound them, you pound them well, until the
stones are pounded well, When we have pounded it, we bring this *tajefnit* (large
dish) where we knead bread, we call it *tajefnit*.
Keep pounding and working on it, pounding and working, bring some charcoal,
put it in the burner, light the fire, and lay on it this *tajefnit,* this dish, leave it
there until it gets really hot and start kneading it with your hands, and you can
see it oil coming out, you knead it like this, like this, and that's it, you bring a
piece of cloth made of linen and press, that's all, and the oil comes out nicely!

Text 4 *ka=n-žni-w*   *z=ze^{y}tun*   **ka=n-žib-u=h**
        PRVB=1-pick\IPFV-PL DEF=N.M.   PRVB=1-bring\IPFV-PL=OBJ.3M.SG
        'We collect the olives, we bring them'
        ka=n-ḥmi-w   *l=fəṛṛan*   *ka=n-ḥmi-w*
        PRVB=1-heat\IPFV-PL DEF=N.M.   PRVB=1-heat\IPFV-PL
        'we heat the oven, we heat it…'
        ka=y-ttə-ḥma   *məzyan bḥal lli ġadi t-ṭiyyəb*
        PRVB=3-heat\IPFV.PASS ADV   CONJ REL FUT 2-cook\IPFV
        'until it is very hot, as if you were going to bake…'
        əl=xobz   **ka=n-məsḥ-u=h**   *məzyan*
        DEF=N.M.   PRVB=1-clean_off\IPFV-PL=OBJ.3M.SG ADV
        '…bread, we clean it off well'
        n-žib-u   *z=zitun*   **n-nəqqi-w=əh**
        1-bring\IPFV-PL   DEF=N.M.   1-clean\IPFV-PL=OBJ.3M.SG
        'we bring the olives, and we clean them'
        wə   **n-ɛəddl-u**   *fi-h*   **n-ɛəddl-u**
        CONJ 1-make\IPFV-PL   PREP-OBL.3M.SG   1-make\IPFV-PL
        'and we work on it, we prepare…'

fi-h   *y-bat*   *l=lila*   *kamla* *y-ibəs*
PREP-OBL.3M.SG 3M.SG-spend_night\IPFV DEF=N.F. ADJ   3M.SG-dry\IPFV
'…it and it spends the whole night (in the oven) in order to make it dry'

t-ḥəss   *bi-h*   *yabəs*
2SG-feel\IPFV PREP-OBL.3M.SG dry\PTCP.ACT.M.SG
'You make sure it is dry'

f-əṣ=ṣbaḥ   *məlli* *t-ḥăll*   *əl=fəṛṛan*
PREP-DEF=N.M. CONJ 2SG-open\IPFV DEF=N.M.
'the next morning, when you open the oven…'

t-ṣib=u   *yabəs*   *t-həzz=u*
2SG-find\IPFV=OBJ.3M.SG dry\PTCP.ACT.SG.M 2-pick_up\IPFV=OBJ.3M.SG
'and you find them dry, you take them out…'

t-žib   *əl=məhraz*   *t-doqq=u*
2SG-bring\IPFV DEF=N.M.   2-pound\IPFV=OBJ.3M.SG
'you bring the mortar and pound them'

t-doqq=u   **məzya(n)** *ḥta*   *yə-tdəqq*
2SG-pound\IPFV=OBJ.3M.SG ADV   CONJ 3M.SG-pound\IPFV.PASS
'you pound it really well, until (its stones) are pounded…'

dik=lə=**εṭam**   *dyal-u*   *məzyan*
DEM.DIST=DEF=N.PL POSS-OBL.3M.SG ADV
'[its stones], well!'

məlli **t-doqq-u**=**h**   *žib-u*   *hadik*   *əəə*
CONJ 2SG-pound\IPFV-PL=OBJ.3M.SG bring\IMP-PL DEM.DIST HESIT
'When you have pounded it, bring this er…'

tažəfnit *baš* *ka=n-εəžn-u*   *ḥna*
N.F   REL PRVB=1-knead\IPFV-PL PRO.IDP.1PL
'*tajefnit* (large dish) where we knead bread, we…'

ka=nqul-u=ha   **tažəfnit**
PRVB=1-say\IPFV-PL=OBJ.3F.SG N.F
'…we call it *tajefnit*.

bqa   **t-doqq**   *u*   *tə-εməl*   *fi=ha*
keep\IMP 2SG-pound\IPFV CONJ 2SG-do\IPFV PREP=OBL.3F.SG
'keep on pounding and working on it'

doqq   *u*   *εməl*   *fi=ha*
pound\IMP CONJ do\IMP PREP=OBL.3F.SG
'pound and work on it.'

žib   *əl=faxəṛ*   **εəml=u**   *f-əl=məžmaṛ*
bring\IMP DEF=N.M. do\IMP=OBJ.3F.SG PREP-DEF=N.M.
'Bring the charcoal, put it in the burner'

šḥəl		*əl=ɛafiya*	*u*	*nəzzəl*	*ɛli-ha*	**dik**	**tažəfnit**
light\IMP		DEF=N.F.	CONJ	lay\IMP	PREP=OBL.3F.SG	DEM.DIST	N.F.

'light the fire, and lay on it this *tajefnit*'

dik=*əl=gəṣɛa*		*xalli=h*		*ḥta*
DEM.DIST=DEF=N.F.		leave\IMP=OBJ.3M.SG		CONJ

'…this dish, leave it there until'

y-sxun		*məzyan*	*u*	*bda*	*ɛəžn=u*
3-heat\IPFV		ADV	CONJ	start\IMP	knead\IMP=OBJ.3M.SG

'it gets really hot and start kneading it.'

b-yəddə=k		*w=əntina*		*ka=t-šuf=u*
PREP-N.M=OBL.2SG		CONJ=PRO.IDP.2SG		PRVB=2-see\IPFV=OBJ.3M.SG

'… with your hands, and you can see…'

ka=y-xərrəž		*z=zit*	*ka=t-ɛəžn=u*
PRVB=3-extract\IPFV		DEF=N.M	PRVB=2-knead\IPFV=OBJ.3M.SG

'…oil coming out, you knead it'

hayda	**hayda**	*ṣafi*	*u*	*t-žib*		**dərra**	*d-əl=ḥayati*
ADV	ADV	ADV	CONJ	2-bringt\IPFV		N.F	POSS-DEF=N.M

'like this, like this, and that's it, you bring a piece of cloth made of linen'

u	*ɛəṣṣəṛ*	*ṣafi*	*u*	*z=zit*
CONJ	press\IMP	ADV	CONJ	DEF=N.F

'and press, that's all, and the oil…'

ka=t-xrəž		*məzyan*
PRVB=3F.SG-come_out\IPFV		ADV

'comes out nicely!'

This technique, with a simple cloth, is not common; more often, a twin-screw wooden oil press is used, like in Taounate (see El Alaoui 2007, Thomas and Caubet 2017, and our own observations with Y. Thomas, L. Clochey and F. El Ghazzaz in April 2014, near Aïn Mediouna). Our informant uses the same technique to make bitter almond oil, because she only makes small quantities at a time.

5.2   Bitter almond oil – zit l-ləwz mərr

Text 5:   *ka-nžibu l-luz bḥal hada u huwa mərr,* **ka-nduqqu-h** *wəlla* **ka-nṭəḥnu-h** *f-Moulinex bḥal lli ġadi nɛəddlu əl-ḥəlwa bda ka-yəṛjəɛ bḥal ṭ-ṭḥin, ṛṭəb, dik-əs-saɛa ka-nžibu əl-kəskas, ka-nɛəmlu* **əl-ma yṭib** *bḥal ila kun-na* **maš nɛəddlu** *əl-couscous, ka-nɛəmlu* **əl-ma yṭibu** *ḥta* **ka-yġliw,** *ka-nɛəmlu* **əd-dərra d-əl-ḥayati** *wəlla ši dərra u ṣafi, əl mohimm* **təqbəṭ** *l-ək ġe l-luz u ṣafi, baš ma yṭeḥ l-ək* **ši** *f-əl-boṛma, tfəṛṛəš-ha f-əl-kəskas, təɛməl* **dik**-*l-luz yətbəxxəṛ ḥta ka-yətbəxxəṛ, ka-yṭeb məzyaaan… məlli ka-təɛ…, hada, ka-tžib-u, k-təbda tɛəṣṣəṛ, ka-txəlli-h ɛla n-naṛ walakin, ma thəbbṭ-u š, ka-txəlli-h, tžib boṭa ṣġira wəlla məžmar*

*wəlla… lli… tžib-u, nəzzəl dik l-boṛma təmma **ḥda-ḵ** u hada.. w-əntina ka-tɛəṣṣəṛ
b-**dik-əd-dərra d-əl-ḥayati** wəlla ka-tɛəddl-u **f-əd-dərra d-əl-ḥayati** u ka-tɛəṣṣəṛ
ġiṛ šwiya, šwiya, šwiya, šwiya, (ḥ)ta…**yhbaṭ z-zit.***

We bring the almonds, like these, but bitter ones (they are bitter), we pound
them, or we grind it in a blender, as if we were going to prepare cakes, and it
begins to look like flour, smooth. Then we take the couscous maker and we
put water to heat as if we were going to prepare couscous. we put water to heat
until it boils, we use a piece of cloth or any cloth, it doesn't matter, as long as
it retains only the almonds that's all, so that it does not fall into the pan, you
spread them out in the colander of the couscous maker, you put the almonds
to steam until they are well steamed, very well cooked; when they er.. what's
it, you bring them, and you begin to squeeze, you leave them on the fire all
the same, don't take them away, you leave them, you bring a small stove, or a
charcoal burner or any… you bring it, and you put this pan (the bottom part of
the couscous maker) there, next to you and the rest, and you go on squeezing
with that cloth, or you do it in the linen cloth and you press it gently, gently,
gently, gently, and the oil comes out

Text 5   *ka=n-žib-u   l=luz   bḥal hada   u*
PRVB=1-bring\IPFV-PL DEF=N.M. ADV DEM.SG.M CONJ
'We bring the almonds, like these, but..'
*huwa   mərr **ka=n-duqq-u=h**   wəlla*
PRO.IDP.3M.SG ADJ PRVB=1-pound\IPFV-PL=OBJ.3M.SG CONJ
'bitter ones (they are bitter), we pound them, or…'
***ka=n-ṭəḥn-u=h**   f-Moulinex bḥal   lli*
PRVB=1-grind\IPFV-PL=OBJ.3M.SG PREP-PN   CONJ REL
'we grind it in a blender, as if …'
ġadi n-ɛəddl-u   əl=ḥəlwa bda   ka=y-ṛjəɛ
FUT 1-make\IPFV-PL DEF=N.F begin\PFV-3M.SG PRVB=3M.SG-become\IPF
'we were going to prepare cakes, and it begins to look…'
bḥal ṭ=ṭḥin   ṛṭəb dik=əs=saɛa   ka=n-žib-u
CONJ DEF=N.M ADJ   DEM.DIST=DEF=N.F PRVB=1-bring\IPFV-PL
'… like flour, smooth. Then we take…'
əl=kəskas   ka=n-ɛəml-u   əl=ma   y-ṭib
DEF=N.M   PRVB=1-do\IPFV-PL DEF-N.M 3M.SG-cook\IPFV
'… the couscous maker and we put water to heat'
bḥal ila   kun-na   maš n-ɛəddl-u   əl=couscous
CONJ CONJ be\PFV-1PL FUT 1-make\IPFV-PL DEF=N.M
'…as if we were going to prepare couscous.'
ka=n-ɛəml-u   əl=ma   y-ṭib-u   ḥta
PRVB=1-do\IPFV-PL DEF-N.M   3M-cook\IPFV-PL CONJ
'we put water to heat until …'

ka=y-ġli-w   *ka=nə-ɛml-u*   *əd=dərra*
PRVB=3M.SG-boil\IPFV-PL   PRVB=1-do\IPFV-PL   DEF=N.F
'it boils, we use a piece of cloth…'[2]

d-əl=ḥayati   *wəlla ši   dərra u   ṣafi əlmohimm*
POSS-DEF=N.M.   CONJ   INDF   N.F   CONJ ADV ADV
'…made of linen, or any cloth, it doesn't matter, as long as…'

tə-qbəṭ   *l=ək   ġe   l=luz*
3F.SG-catch\IPFV PREP-OBL.2SG ADV DEF-N.M
'it retains only the almonds'

*u   ṣafi baš   ma   y-ṭeḥ   l=ək   **ši***
CONJ ADV CONJ NEG1 3M-fall\IPFV PREP-OBL.2 NEG2
'that's all, so that it does not fall…'

f-əl=boṛma   t-fəṛṛəš=ha   f-əl=kəskas
PREP-DEF=N.F 2SG-spread_out\IPFV=OBJ.3F.SG PREP-DEF=N.M
'…into the pan, you spread them out in the colander of the couscous maker…'

*tə-ɛməl   **dik=l=luz**   yə-tbəxxəṛ*
2SG-do\IPFV DIST=DEF=N.M 3M.SG-steam\IPFV.PASS
'you put the almonds to steam,'[2]

*ḥta   ka=yə-tbəxxəṛ   **ka=y-ṭeb**   məzyan*
CONJ PRVB=3M.SG-steam\IPFV.PASS PRVB=3M.SG-cook\IPFV ADV
'…until they are well steamed, very well cooked'

məlli ka=t-əə   hada ka=t-žib=u
CONJ PRVB=3F.SG-HESIT DEM PRVB=2SG-bring\IPFV=OBJ.3M.SG
'when they er.. what's it, you bring them'

ka=tə-bda   t-ɛəṣṣəṛ   ka=t-xəlli=h
PRVB=2-begin\IPFV 2-press\IPFV PRVB=2-leave\IPFV=OBJ.3M.SG
'and you begin to squeeze, you leave them'

ɛla   n=naṛ   walakin ma   t-həbbṭ=u   š
PREP DEF=N.F. CONJ   NEG1 2SG-take_down\IPFV=PRO.3M.SG NEG2
'.. on the fire all the same, don't take them away'

ka=t-xəlli=h   t-žib   boṭa ṣġira wəlla məžmar
PRVB=2-leave\IPFV=PRO.3M.SG 2-bring\IPFV N.F   ADJ.F CONJ   N.M
'you leave them, you bring a small stove, or a charcoal burner'

wəlla lli   t-žib=u   nəzzəl   dik=l=boṛma
CONJ   REL 2SG-bring\IPFV=OBJ.3M.SG put\IMP DEM.DIST=DEF=N.F
'or any, you bring it, and you put this pan (the bottom part of the couscous maker)'

*təmma **ḥda-ḵ**   u   hada w-**əntina***
ADV.LOC PREP-OBL.2 CONJ DEM CONJ-PRO.IDP.2
'there, next to you and the rest, and you'

ka=t-ɛəṣṣəṛ   *b-**dik**=əd=dərra*   *d-əl=ḥayati*   *wəlla*
PRVB=2SG-press\IPFV   PREP-DIST.DEM=DEF=N.F.   POSS-DEF=N.M   CONJ

'(you) go on squeezing with that cloth, or'

ka=t-ɛəddl=u   *f-əd=dərra*   *d-əl=ḥayati*
PRVB=2-make\IPFV=OBJ.3M.SG   PREP-DEF=N.F.   POSS-DEF=N.M.

'you do it in the linen cloth'

u   *ka=t-ɛəṣṣəṛ*   *ġiṛ*   *b-šwiya*
CONJ   PRVB=2-press\IPFV   ADV   PREP-ADV

'and you press it gently…'

šwiya   *šwiya*   *šwiya*   *u*   *ka=y-hbəṭ*   *z=zit*
ADV   ADV   ADV   CONJ   PRVB=3M.SG-go_down\IPFV   DEF=N.F

'gently, gently, gently, and the oil comes out'

These texts give us insights into techniques still used in the Rif; we will now try and analyse the linguistic phenomena due to language contacts.

6.   Arabic and Berber in contact

6.1   Language contact

The contact between Berber and Arabic dates back to the arrival of Arabic in the region, in the 7th century, and the cross-influences are deep and intricate. The Prehilali Arabic dialects of North Africa all bear traces of these contacts. But among them, what William Marçais called "les parlers villageois", and later, Colin for Morocco, "les parlers montagnards", the oldest and most innovative Arabic dialects in the Maghrib are the descendants of the first layers of Arabization (see Marçais & Guîga 1925 and Colin 1937). Berber and Arabic share a long history of contact and most studies have shown how much Berber has borrowed from Arabic, especially on the lexical side. The situation presented here is a result of a historical situation, combined with present-day regular contacts.

Kossman (2017b) in an article on Berber-Arabic language contact writes:

> Since the start of the Islamic conquest of the Maghreb in the 7th century CE, Berber and Arabic have been in continual contact. This has led to large-scale mutual influence. The sociolinguistic setting of this influence is not the same, though; Arabic influence on Berber is found in a situation of language maintenance with widespread bilingualism, while Berber influence on Arabic is no doubt to a large degree due to language shift by Berber speakers to Arabic.

In the case of Msek, we have seen that the child is led to learn Berber in order to communicate with other children at school, so it is an unusual situation. Kossman's study is very detailed, analyzing sound systems, morphology, syntax and lexicon:

> Arabic influence on Berber is the result of a long history of coexistence, with high degrees of bilingualism on the part of Berber speakers. (…) On the other side of the coin, Berber influence on Arabic is no doubt to a large degree due to language shift from Berber to Arabic (on which see, among others, Lévy 1998), as minority Arabic-speaking groups gradually assimilated more and more Berber speakers and groups.

As for the data we collected, we will examine some linguistic features that can be linked to contact and have been studied by dialectologists for the past century:[11] on the phonetic level: vowels and diphthongs, spirantization, weakening of certain fragile phonemes; an interesting question on the status of *u/w* - vowel or consonant – and its implications on verbal morphology; morphosyntax (preverb of the imperfect, future particle, demonstratives…), and changes in agreement in gender and number due to contact; a rare negation particle *bu* present in both languages; and finally a few specific lexical features.

6.2   Phonetics: Remarks on vowels and diphthongs

–   In **Taounil**, the vowels have a particular colour in F.'s speech: the /ə/ is realized [œ] *l-lḥœm* "meat", in a pharyngeal context, or [ă] near a glottal, *ka-yakl-uw-ăh* "they eat it".
–   There are some diphthongs – which are found among women, according to Vicente (2000: p. 34, 2005: p. 112); for **Taounil**: *ət-tăwm, əl-băyṭat, əl-lăwz* "garlic, eggs, almonds"; for **Msek**: *l-ləuz, z-ze͏ʸtun* "almonds, olives".
–   Clearly linked to F.'s Berber "accent", the nasal realization of final *-a* in pausal position: *f-əl-mã*, "in water", *mən təmmã*, "from there":

(1)   əl=ḥaža   r=rumiya   ka=y-ɛəml-u   l-a   əd=dwã
       DEF=N.F. DEF=ADJ.F. PRVB=3-do\IPFV-PL PREP-OBL.3F.SG DEF=N.M
       'the imported plant (thing), they give it treatment (medicine),
       əl=bəldiya   ma   ka=y-ɛəml-u   l-a   bu[12] d= dwã
       DEF=N.F   NEG1 PRVB=3-do\IPFV-PL PREP-OBL.3F.SG NEG2 DEF=N.M
       the local one doesn't need any treatment'

---

11. For a summary of these points of variation in Arabic dialects, see David Cohen's questionnaire (Cohen & Caubet 2000), and the one I proposed for North Africa (Caubet 2002).

12. For negation and the use of *bu*, see below 6.8.

6.3   Phonetic remarks on consonants and spirantization

Some features are also found in the Jbala dialects, like the unvoiced realization of *q* and of *ḍ*.

- /q/ is realized [q]: *fuq* "on", *səlqu* "they boiled", except in the word *waxt* for *wəqt* "time" in **Taounil**; see also the compounds formed with *wəqt* (Colin:1921: p. 43, for North Taza: *fiwax* (<*fi ayy weqt) "when", *dūx* (<*da l-weqt) "now").
- /ʕ/ in realized [ḥ], in **Msek**: *šḥəl* (šɛəl) "light".
- /ḍ/ is realized [ṭ] *əl-băyṭat* "the eggs" in **Taounil**, as is often the case in Jbala dialects (see Colin 1921: p. 40 for North Taza and for Ceuta, Vicente 2005: p. 115). In Msek: *əl-ɛṭam* "the bones".
- In **Taounil**, /ž/ is realized [ž], *ṭanža* "Tangiers", *žit* "I came, I am here", except when it is geminated, *ž + ž > ǧ: (ǧ)ǧdad* "chicken", *(ǧ)ǧiran* "the neighbours", like in North Taza (see Colin 1921: p. 39). It is the same in **Msek**: *ka-n-žib-u-h* "we bring it", *ka-t-ɛəžn-u* "you knead it", *əl-məžmaṛ* "the charcoal burner". I found no example of *ž + ž > ǧ*.
- **Spirantization** is not generalized, but can be found in **Taounil** with /b/, /k/ and /d/: *l-lḇən* "buttermilk", *haḏi* "this one (fem.)", *hayḏa* "like this", although otherwise realized [b] and [d]: *bḥal, bġiti, hadak; hayda* "like, you want, that one (masc.), like this". In **Msek** (*msek̲*): *ḥḏa-k̲* "next to you", but *hadik* "that one (fem.)", *bda* "he began", *hayda* "like this", *dərra* "handkerchief", *bḥal* "like", *f-əl-boṛma* "in the pan".

6.4   The weakening of fragile phonemes /h, l, n/

There are many cases of elision in Prehilali dialects: apocope, apheresis; and the 3rd person pronouns, *-ha* (fem.) and *-hum* (plural), where the *h* is often elided.

- **Truncation:** in **Msek**, there is an *apocope*: *məzya(n)* "well", and the indefinite article *wāḥəd-əl-bənt* "a girl", which is reduced to *wäḥ* (see Maghdad 1993): *wäḥ*-l-moḏaʕ "a place", *wäḥ nnhaṛ* "one day". I found a case of apheresis in 2014: *f-aḥəd-lə-žəbbānya* "in a large bowl". The article is reduced to *ḥa* in North Taza (see Colin 1921: p. 30): *ḥannhāṛ* "one day".
- **Elision of the *h* in the affix pronouns** *-ha* and *-hum; elision of the final l of dyal*

In **Msek**, the *h* of the affix pronouns is often elided:

(2)   *ma*   *ka=n-dir=***a**   *ši*   *bəzzaf*
      NEG1   PRVB=1SG.do\IPFV=OBJ.3F.SG   NEG2   ADV
      'I don't make it much anymore'

(3) *nə-mši*   *n-žib*=**a**   *l-**a***
     1SG-go\IPFV 1SG-bring\IPFV=OBJ.3F.SG PREP-OBL.3F.SG
     'I'll go and ge it.'

(4) *ka=nə-ɛṭi-w=ha*   **l-um**
     PRVB=1-give\IPFV-PL=OBJ.3F.SG PREP-OBL.3PL
     'We give it to them.'

(5) *ka=y-žīb(u)=**um***   *l-na*   *l=fəllāḥa*
     PRVB=3-bring\IPFV-PL=OBJ.3PL PREP-OBL.1PL DEF=N.PL
     'The peasants bring them to us.'

But it is also present:

(6) *ka=nə-ɛṭi-w=**ha***   *l-um*
     PRVB=1-give\IPFV-PL=OBJ.3F.SG PREP-OBL.3PL
     'We give it to them.'

(7) *ka=yə-ɛṭī-w=**ha***   *l-na*
     PRVB=3-give\IPFV-PL=OBJ.3F.SG PREP-OBL.1PL
     'They give it to us.'

There is also a very strange phenomenon which was described for the first time by Maghdad for **Msek**, and commented by her supervisor, Simon Lévy (1998: p. 12 note 6): **the elision of the final 'l' of possession particle *dyal*:**

> (…) the occlusives /b/, /t/, /d/, /ḍ/, /k/ are slightly fricative /b/, /t/, /d/ like in *Tarifit*; the liquid /l/, realized /r/ /ž/ in *Tarifit,* is muted in the studied dialect: *dyäl-i > dyäy* (my/mine).[13]

The "l" of the possessive *dyal* is assimilated or elided; the complete paradigm was given by Maghdad 1993: *dyäy, dyäk, dyänna,* but for the 3rd person masculine, *dyäh;* this is because the fall of "l" entails the presence of the "h", by compensation: *dyäl-i > dyä-y* 'my, mine' but, *dyäl-o > dyä-h* 'his', *dyäl-hom > dyä-hom* 'their(s)'.

    In our 2014 data, there is a hesitation, but the pronunciation is so fast and lax that it is sometimes difficult to hear if the "h" is really present or whether it is just a diphthong: *ən-nəṣṣ dya~um* "heir half"; *lḥăqq dya~hum* "their share", *dya~ha* "hers".

    In **Taounil**, we found an elision of the final l of *bḥal* "like": *bḥa hada* "like this", and the muting of the *h* in affix personal pronouns is common:

---

13. (…) les occlusives /b/, /t/, /d/, /ḍ/, /k/ sont réalisées légèrement fricatives /b̠/, /t̠/, /d̠/, /ḍ̠/, /k̠/ comme en tarifit; la liquide /l/ réalisée /r/ /ž/ en tarifit – s'amuït dans le parler étudié ». *dyäl-i > dyäy* (mon/à moi).

(8) *l=εarbiya   tεəlləm-t=**a**   ġe   hna*
     DEF=N.F.   learn\PFV-1SG=OBJ.3F.SG   CONJ   ADV
     'Arabic, I learnt it just here.'

(9) *εənd=**om**   bəzzaf*
     PREP=OBL.3PL   ADV
     'They have a lot.'

(10) *ka=tə-nbət   b-waḥd=**a***
      PRVB=3F.SG-grow\IPFV   PREP-NUM=OBL.3F.SG
      'It grows by itself.'

(11) *ṭ=ṭiba   dyal-**a**   f-əš=škəl*
      DEF=N.F.   POSS-OBL.3F.SG   PREP-DEF=N.M
      'Its taste is different.'

6.5   Vowel or consonant: u/w? Implications on verbal morphology, the sense of an evolution

In Taounil, when F. was describing the way she prepared *tasukkant* or *zembu*, a certain rhythm was given to her narration, by the series of plural verbal forms expressing the habitual. They were all imperfect plurals with a 3rd p. M.SG affix pronoun (see texts 2 and 3):

> *ka-nsəlquw-ăh* "we boil it", *ka-nṭəyybuw-ăh* "we cook it", *ka-nεəmluw-ăh* "we do it", *ka-nεəṣṣṛuw-ăh* "we press it", *ka-nṭayybuw-ăh* "we cook it", *ka-nfowwṛuw-ăh* "we steam it".

The 3rd p. M.SG affix pronoun has two forms in M.A., depending of the ending of the word: *-h*, if the word ends in a vowel, and *-u*, if it ends in a consonant: *εli-***h** "on it/him"; *l-***u** "to him". In F.'s speech we find **an innovation**: *ka-n-εəml-u-h > ka-n-εəml-u***w**-***ăh**, as illustrated in the list above. Over ten verb forms, only one was regular, *ka-yεəmlu-***h.**

  This form is quite frequent in Prehilali dialects, but only with *defective verbs*; in F.'s case, it is found with *regular verbs*. We'll try and analyse **the sense of an evolution**.

–   ***For defective verbs only in the North of Morocco, Djidjelli and Tlemcen***

This construction happens with ***defective verbs*** when the affix is added to a vocalic ending, thus producing a diphthong: *nsa-w + h > nsaw-ăh* "they forgot" "they forgot it".

  Msek: with S., the *-uw-ăh* ending occurs only with ***defective verbs***: *n-nəqqi-w-əh* "we'll clean it"; *nxəlliw-əh yxməṛ* "we let it rise (the dough)", whereas regular verbs

have the regular construction: *ka-nžibu-h* "we bring it"; *ka-n-məsḥ-u-h* "we wipe it clean"*; ka-nduqqu-h* "we pound it"; *ka-nṭəḥnu-h* "we grind it".

These forms have indeed been described for *defective verbs* for **Tlemcen** (W. Marçais 1902: p. 130), **Tangiers** (W. Marçais 1911), **North Taza** (Colin 1921: p. 71), **Djidjelli** (P. Marçais 1956: p. 441), **Fez** (Caubet 1993: p. 161), **Chaouen** (Natividad 1998: p. 117), and Moscoso (2003: p. 162).

Tlemcen: W. Marçais (1902: p. 130) mentions it for *defective verbs* only, and he describes it as a *"consonantization"* of *u* in *w*:

> Il importe enfin de noter la singulière façon dont les pluriels en *âu, îw*, provenant de parfaits et de futurs de verbes défectueux se comportent avec les affixes personnels. (…) les affixes de la 2ᵉ pers. sing et 3ᵉ ms. sing. sonnent *ök, åk* et *öh, åh*, et *u* se consonantise devant eux en *w*. L'on a ainsi de *qrâu*, ils ont recité, *qrâwöh*, ils l'ont recité ….

I found an example in his texts (1902: p. 264 l. 1): *yenfiwöh* "they banish him"; with ordinary verbs, the construction is a simple affixation of pronoun *-k* (1902: p. 268 l. 42): *nḥebbrûk* "we inform you".

Tangiers: W. Marçais (1911) does not give a description of the dialect, but I was able to find one occurrence in the texts (1911: p. 57), on a *defective verb* (*w* is transcribed *u̯*): *εómᵐṛom ma ḳệʾḫ̣ȧṭṭệu̯ah* "ils n'y manquent jamais" (translation by Marçais 1911: p. 162 "they never fail to do it"), with verb *ḫȧṭṭȧ* (1911: p. 282) "négliger, manquer à l'accomplissement d'un devoir – to fail to accomplish/do".

North Taza: Colin (1921: p. 71) also describes it for *defective verbs*:

> Suffixés à l'une des personnes du pluriel des verbes défectueux, les affixes de la 2e pers. comm. sing et de la 3e pers. masc. sing, deviennent respectivement: *-ök* et *-ah*, *-öh*; le و de la terminaison se consonantise. *wuṣṣâwah* ils l'ont arrangé; *nəbyîwök* nous t'aimons; *rmîwöh* jetez-le.

Djidjelli: P. Marçais (1956: p. 441) mentions it only for *defective verbs*:

> Quand le thème verbal comporte une syllable diphtonguée, *āu, īu* (…) L'adjonction de l'affixe de la 2ᵉ personne du singulier, et bien plus encore celle de l'affixe de la 3ᵉ du masculine, font apparaître *des ensembles de constitution phonique insolite* don't il faut tenter de rendre compte: *nsẹ̣u + ĕk = nsẹ̣wĕk; nsẹ̣u + ĕh = nsẹ̣wĕh*; (…)

He analyses it as due to the ambiguity of the second element of the diphthong, *u/w*:

> (…) on mettra en cause la nature ambiguë du u/w second élément de ces diphtongues: senti ou traité comme voyelle, il aboutit à *nsẹ̣uk, nsẹ̣uh* (…) senti et traité comme consonne, il doit aboutir à: *nsẹ̣wĕk; *nsẹ̣wu* (…) *nsẹ̣u + k* étant *nsẹ̣wĕk, nsẹ̣u + h* a été *nsẹ̣wĕh*; (…)

He then quotes an example from Tangiers (W. Marçais 1911: p. 23 l. 5), where the same construction applies to an active participle: *ɛāmlînǎh*,[14] which he explains as being influenced by a combination of two constructs: *ɛāmlînĕk* and *yǎɛṭîwǎh* – *in* + *h* being treated like *īu* + *h*. He also mentions this construction in his *Esquisse* (P. Marçais 1977: p. 193).

Fez: I had also noted this form for *defective verbs* (Caubet 1993 tome 1: p. 161): *nsāw-ək* "ils t'ont oublié(e); *nsāw-əh* "ils l'ont oublié".

Chaouen: Natividad mentions it for *defective verbs* (1998: p. 117):

> Les verbes défectifs au pluriel de l'inaccompli intercalant une voyelle brève /ə/ entre la désinence du pluriel et les pronoms suffixes des 2ᵉ et 3ᵉ personnes du singulier: /nā-nġəllīwəh/ "nous le bouillons"; /nəʕʕīwəh/ "nous le faisons"; *nəʁmīwəḵ* "nous allons t'abandonner".[15]

Moscoso (2003: pp. 162–163) gives examples for *defective verbs*: *n-nās lā-yǎɛṭīwǎh* "people give it normally"; he adds that in rare cases, it can be extended to *regular verbs* and he gives one example: *nā-nṭǎhhʁuwǎh* "we circumcise him", adding that the norm would be *-ūh*.

The Chaouen situation will lead us to consider this innovative construction in dialects where the *-uw-ǎh* ending is frequent for *regular verbs*: Anjra and Ceuta and our new data, Taounil.

– *Innovation, the new construct for all types of verbs: Anjra, Ceuta, Taounil*

Anjra: A. Vicente first reported the extension of the construct when she described the dialect of Anjra (Vicente 2000: p. 138):

> (…) en el dialecto de Anjra, esta manera de realizar el sufijo se ha extendido a todos los demás tipos de verbos (In the dialect of Anjra, this realization of the suffix has extended to all the other types of verbs.

She gives a number of examples: *nsǎmlūwǎh* "we do it"; *nǧîbūwǎh* "we'll bring it"; *nxǎslūwǎḵ* "we'll wash you"; *yṭǎyybūwǎh* "they cook it".

Ceuta: Vicente also found this construction in her description of Ceuti Arabic (2005: p. 153); she describes the phenomenon as: "la présence dans les formes verbales d'une voyelle brève /ǎ/ entre les désinances verbales du pluriel et les suffixes *-h* et *-k*. Ce trait existe dans tous les types de verbes." She gives the following examples: *nāklūwǎh* "we eat it", *nsĭyybūwǎh* "we throw it ways".

---

14. Instead of *ɛāmlînu*.

15. We use the authors' original transcriptions.

Taounil: F. spent 15 years in Ceuta in a Tarifit speaking family; did she pick it up in Ceuta, or is it present in the area? We would need to enquire more. A very peculiar rhythm is given to text 2: *ka-nsǝlquw-ăh, ka-nṭǝyybuw-ăh, ka-nɛǝmluw-ăh, ka-nɛǝṣṣṛuw-ăh, ka-nfowwṛuw-ăh,* etc.

– *The sense of an evolution: regularization*

We saw P. Marçais's hypothesis (1956: p. 441) about the ambiguity of the status of the phoneme /*u*/*w*/ to explain the sense of this evolution. Vincente (2000: p. 138) noted[16] "the presence in the verbal forms of a short vowel /*ă*/ between the plural verb ending and the suffixes *-h* and *-k*".

In fact, this diphthong *adds a syllable* to the word and gives it more weight and depth:

> *kayɛǝmlu-h* is pronounced as *three syllables*: *kay – ɛǝm – luh.*
> *kayɛǝmluw-ăh*, as *four syllables*: *kay – ɛǝm – lu – wăh;*

Expressivity uses the lengthening of words, by the addition of suffixes for example (which also adds syllables): *hna* and *hna-ya* "here", *hna-k* "there"; *ḥna* and *ḥna-ya* "we, us"; *hayda* and *haydaya* "like this"; this new verbal form probably gives the same effect.

The need for regularization of a paradigm can lead to evolution. We will explain the process as follows: **defective verbs** evolve first: from *nsa-u-k / nsa-u-h* to *nsa-w-ǝk / nsa-w-ăh* "they forgot you/him"; from *nnǝqqi-u-h* to *nnǝqqi-w-ǝh* "we will clean it"; once this is established, the **regular verbs** can align on the paradigm, like in Anjra, Ceuta and, as we just discovered, Taounil.

6.6   Morphosyntax

We'll list some features that characterize these two dialects: the preverb of the imperfect, the demonstrative adjectives and the future particle.

6.6.1   *The preverb of the imperfect is exclusively ka-*
Contrary to a number of Jbala dialects that have a variety of other preverbs, *ʾa-*, *a-*, *la-/na-* or *da-*,[17] both varieties studied here have *ka-* – the koinic preverb – exclusively: *ka-y-ɛǝml-u* etc.

---

16. My translation.

17. See Colin (1921), Lévy-Provençal (1922), Natividad (1998), Vicente (2000), Moscoso (2003), Caubet (2017).

6.6.2   The demonstrative adjectives are invariable

Had (proximal) is invariable in M.A., but in Msek and Taounil, *dik* (distal feminine form) also is; it is the same in Anjra (Vicente 2000: p. 139). Colin notes *dak* invariable for North Taza (Colin 1921: p. 71): fem. *dik-əs-saɛa* "that time", plur. *dik-lə-ɛṭam* "those bones", masc. (collective) *dik-l-luz* "those almonds". The demonstrative adverb is *hayda* "like this".

6.6.3   From collective to plural

The term *əl-bayḍ / əl-beḍ* "(the) eggs" is a collective in M.A. In F's speech in **Taounil**, it becomes countable and takes a plural: *əl-băyṭ-at*; *ḍ > ṭ* and an external plural is formed with the suffix *-at*.

6.6.4   Possession: Double construction

In **Msek**, I found a double construction used with kinship names, which had been noted for Djidjelli, where it is generalized (P. Marçais 1956: p. 413, 421): the possessor is marked doubly, via the possessive pronoun which is affixed on the kinship term, and through the analytic construction, with possessive particle *d*: *ḥnaya, bba-h d Moḥammǝd, ɛǝnd-u l-arḍ bǝzzaf*:

(12)   *ḥnaya*   *bba=h*   *d*   *Moḥammǝd…*   *..ɛǝnd=u*
      PRO.IDP.1PL   N.M=OBL.3M.SG   POSS   PN   PREP=OBL.3M.SG
      l-arḍ   *bǝzzaf*
      DEF=N.F.   ADV
      'As for us, Mohammed, his father, he has a lot of land' [lit. his father of M.]

6.6.5   The future particle

I only found future particles in S.'s data in **Msek**; she uses *ġadi,* the usual particle for central Morocco, and several occurrences of the more Prehilali particle, *maš:*

(13)   *bḥal*   *lli*   **ġadi**   *t-ṭiyyəb*   *əl=xobz*
      CONJ   REL   FUT   2SG-cook\IPFV   DEF=N.M.
      'As if you were going to cook bread.'

(14)   *bḥal*   *lli*   **ġadi**   *n-ɛǝddl-u*   *əl=ḥǝlwa*
      CONJ   REL   FUT   1-make\IPFV-PL   DEF=N.F.
      'As if we were going to make cakes.'

(15)   *bḥal*   *ila*   *kun-na*   **maš**   *n-ɛǝddl-u*   *əl=couscous*
      CONJ   CONJ   be\PFV.1PL   FUT   1-make\IPFV-PL   DEF=N.M
      'As if we were going to make couscous.'

(16)   *əṣ=ṣǝḥfa*   *f-aš*   **maš**   *n-ɛǝžn-u*
      DEF=N.F.   PREP-REL   FUT   1-knead\IPFV-PL
      '…the dish in which we are going to knead (the dough)'

6.7   Changes in agreement in gender and number due to language contact

We will examine the influence of both languages on the agreement in gender and in number. Some changes in gender or number have been lexicalised, others are due to an accidental confusion.

– In **Taounil**, F., a second language speaker often hesitates in **gender** agreement; in text 3, she gets confused in the agreement of *zembu* (masc.) and *dšiša* (fem.). This is flagrant in the first part of text 3: *ha **huwa** əd-dšiša, ha əd-dšiša! zəmbu ha **hiya**, ha **huwa**! hadi əd-dšiša u **hadi** zəmbu!*

– | *ha* | *huwa* | *əd=dšiša* | *ha* | *əd=dšiša* | *zəmbu* | *ha* | *hiya* |
  |---|---|---|---|---|---|---|---|
  | PRST | PRO.IDP.3M.SG | DEF=N.F | PRST | DEF=N.F | N.M | PRST | PRO.IDP.3F.SG |

ha	*huwa*	*hadi*	*əd=dšiša*	*u*	*hadi*	*zəmbu*
PRST	PRO.IDP.3M.SG	PROX.F.SG	DEF=N.F	CONJ	PROX.F.SG	N.M

  'This (masc.) is 'dchicha', here is 'dchicha'!' Zembu', here it (fem) is! Here it is (masc.); this is 'dchicha', this (fem.) is 'zembu'!'

It also happens in text 3, where *zəmbu* (masc.) has feminine agreement: ***haḍi** ka-yakl-u-**ha** hayḍa etc.* "this one (fem.), they eat it (fem.) like this"; whereas *dšiša* (fem.) has masculine agreement: *u **hadak** ka-n-ṭayyb-uw-**ăh** etc.* "that one (masc.), we cook it (masc.)…".

Similarly in text 2, *tasukkant*, which is a feminine word in Berber, agrees systematically in the masculine in text 2: *ka-n-səlq-uw-**ăh** f-əl-mã̃,* etc. "we boil it (masc.) in water…" We saw in 4.1 the ambiguity about the origin of the term; the agreement may be due to the word in M.A., *səkkum,* which is masculine… But in text 1, it has feminine agreement. This shows F.'s confusion on the question of agreements.

– For **number**, some words that are usually singular in M.A. have become plural in this dialect under the influence of Berber; for example ***əl-ma*** "water" (*aman,* pl. in Berber) or ***l-qmăḥ*** "wheat", which has plural agreement, due to Berber influence:[18]*b-əl-ma sxun-in* (adj. plural); see Colin 1921: p. 58; Salem Chaker confirms that "*irden*, le blé, terme pan-berbère, est bien un pluriel et commande un accord pluriel (vrai aussi pour l'orge, *timẓin, tumẓin*)."

Taounil: it has plural agreement with an adjective: *b-əl-ma sxun-in* "with hot water".

Msek: S. starts with a singular agreement: *kanɛəmlu əl-ma yṭib* "we put water to boil (sg.)"; but she immediately switches to the plural (text 6) on bitter almond oil: *kanɛəmlu əl-ma yṭibu, ḥta **kayġliw*** "we put water to boil (pl.), until it (pl.) boils. In a long passage about the making of flour, ***l-qmăḥ*** has plural agreement:

---

18. Personal communication from Salem Chaker (Sept. 2014).

*kanžibu dīk-əl-**qmăḥ**, kanəqqiw-**hum** u nəddiw-**hum** l-ər-rḥa, nṭəḥnu-**hum** fə-r-rḥa…*

'We bring that wheat, we clean it (pl.), and we take it (pl.) to the mill, we grind it (pl.) in the mill…'

6.8   Negation: Conservatism and innovation, locator **bu**

6.8.1

A first remark is that in **the discontinuous morpheme *ma…š*,** the second element (< *šay "thing") is often found in a fuller form, *ma…ši, ma…šay*, without rendering the negation more insistant, as would be the case in central Morocco (see Caubet 1993 tome II: p. 68).

In **Msek**:

(18)   *u   ila   **ma**   ɛməl-ti-ha   šay*
        CONJ CONJ NEG1 do\PFV-2F.SG=OBJ.3F.SG   NEG2
        'And, if you don't do this…'
        *baš   **ma**   y-ṭeḥ   l-ək   ši*
        CONJ NEG1 3M.SG-fall\IPFV   PREP-OBL.2SG NEG2
        'that's all, so that it does not fall…'

(19)   ***ma**   ka-n-dir-a   ši   bəzzaf*
        NEG1 PRVB=1.do\IPFV=OBL.3F.SG   NEG2 ADV
        'I don't make it much anymore (almond oil)'

But there are also occurrences of *ma…š*:

(20)   ***ma**   thəbbṭ-u   š*
        NEG1 2-take_down\IPFV=OBL.3M.SG   NEG2
        '… on the fire all the same, don't take them away'

(21)   *ḥna   **ma**   ka-n-həḍr-u   š   š-šəlḥa*
        PRO.IDP.1PL NEG1 PRVB=1-speak\IPFV-PL   NEG2 DEF=N.F.
        'We don't speak Berber here!'

In **Taounil,** we don't have enough examples to draw any conclusions and we would need more data, but the only two forms we found have a reduced form:

(22)   ***ma**   ka=y-fəhm-u   š   məzyan bḥal ḥnaya*
        NEG1 PRVB=3-understand\IPFV-PL   NEG2 ADV   CONJ PRO.IDP.1PL
        They (Beni Iṭṭeft) don't understand (Berber) well, like we do (Arabic)

(23)   *huma*   *Bni Bufraḥ*   **ma**   *y-qədṛ-u*   *š*   *y-həḍṛ-u*
       PRO.IDP.3PL   Beni Boufrah   NEG1   3-can\IPFV-PL   NEG2   3-speak\IPFV-PL
       huma   *š=šəlḥa*
       PRO.IDP.3PL   DEF=N.F.
       'For their part, Beni Boufrah, they cannot speak, them, Berber…'

6.8.2

Another remark concerns **the presence of a second element *bu*** for strong negation in Taounil: *ma…bu* "not…at all". It is also present in Maghdad (1993)[19] *mä käyn **b̠u ḍḍṛa*** "there was no corn at all", with a spirantization of the *b* in *b̠u* and an article, *ḍ=ḍṛa*. In February 2016, I asked Amal Maghdad to give me an occurrence with a non-assimilating consonant to make sure the article was present: *mä käyn **bu l-ḥlib*** "there is no milk at all".

I had described this negation at length for Fez (Caubet 1983: pp. 172–176); at the time I had not found it described in any previous publication and it was really *novel*. It struck me when I found it in Maghdad's (1993) data in 2012, and later, when I heard F. using it in 2014 – note that *bu* bears a stress:

(24)   *əl=ḥaža*   *r=rumiya*   *ka=y-ɛəml-u*   *l=a*   *əd=dwæ̃*
       DEF=N.F.   DEF=ADJ.F.   PRVB=3-do\IPFV-PL   PREP=OBL.3F.SG   DEF=N.M
       əl=bəldiya   **ma**   *ka=y-ɛəml-u*   *l-a*   **ˈbu**[20]   **d=dwæ̃**
       DEF=N.F   NEG1   PRVB=3-do\IPFV-PL   PREP-PRO.3F.SG   NEG2   DEF=N.M
       'The imported plant (thing), they give it treatment (medicine), the local one doesn't need any treatment.'

In my 1978–1980 Fez data (see Caubet 1983: pp. 172–176), I first heard it from the mother, who was around 60 at the end of the 70's, but the whole family (of Riffian origin) used it, although many Moroccans seemed surprised when they heard it, or found it too "rural".

In 1983, I worked in the enunciative framework, with Antoine Culioli, and I still find the analysis pertinent: *bu* (< *father) acts as a locator (see below its usages), and an example like *ma kain **ˈbu ḥlib*** (with no article in Caubet 1983: p. 172, and with a stress) "there is no milk at all", could be glossed as: "there is a relation between *bu* and *ḥlib*; when negating the locator (*bu*), one negates the localisation of the object, and thus, its existence, leading to a strong negation bearing no exception".

The difference with Taounil and Msek is that there is **an article** before the negated term, whereas in my Fez data, there is clearly **no article**. I checked again, thirty years later in February 2016 with one of my informants to make sure, with

---

19. See Maghdad (1993: p. 23), Text II, and Caubet (2016, 2017).

20. For *bu* in negation, see Caubet (1983: p. 172–174) and Caubet (1993 tome II: p. 294–295).

nouns starting with lunar consonants. *Bu* is often used in an answer to a previous utterance or a question and comes as a strong denial:

(25) – *waš ma   ɛənd=ək   əl=mæednus?*
      Q   NEG1 PREP=OBL.2SG DEF=N.M
      – *ma ɛənd=i   'bu   mædnus*
      NEG1 PREP=OBL.1SG NEG2 N.M
      '– Don't you have any parsley?'
      '– No I don't have any parsley/the slightest sprig of parsley.'

(26) – *žibi   ġṭaṛ   mən   təmm*
      bring\IMP.F N.M   PREP ADV
      – *ma kayn 'bu   ġṭaṛ*
      NEG1 EXS   NEG2 N.M.
      '– Bring a plate from there!'
      '– There is no (not a single) plate here!'

My informant spontaneously gave me two synonyms of strong negation and, as noted in Caubet (1983 and 1993), they differ in nominal determination: *bu* is followed by the bare noun, whereas *š* needs a construction with a determined substantive:[21]

(27) *ma   ɛənd=i   'bu   luz*
      NEG1 PREP=OBL.1SG NEG2 N.M

(28) *ma   ɛənd=i   š   l=luz*
      NEG1 PREP-OBL.1SG NEG2 DEF=N.M
      'I don't have any almonds at all.'

Utterance (27) can be glossed as: "there is a relation of location between *bu* and *luz*, when negating the locator (*bu*), you negate the localisation of the object, and thus, its existence; as far as *luz* is concerned, it does not exist". (See Caubet 1983: p. 176.)

Utterance (28) could be glossed as: "whichever way you look at it, there is no *šay/š* "thing"/"trace" (not the slightest trace) of a validation of the relation 'me-having almonds'; the predicative relation is negated". (See Caubet 1983: p. 89.)

The negation of the location or of the complete predicative relation produces an excluding negation.

What is common to the Fez and Al Hoceima examples is that *ma…bu* marks this strong excluding negation, whereas recent work by Mena Lafkioui about the "Moroccan Arabic variety of Oujda (MAO)" shows a different functioning. In her

---

21. See Caubet (1983: pp. 89–90) for the use of article *əl*.

article (Lafkioui 2013), she claims that *bu* is only used in Oujda Arabic, and she gives it a Berber origin:

> With respect to this negation system, the Moroccan Arabic variety of Oujda distinguishes a new discontinuous marker, *ma ___ bu,* whose second element is borrowed from Tarifit, which is the only Berber language where this morpheme occurs. (Lafkioui 2013: p. 51)

About the Berber etymology, she tries to give *bu* a negative meaning; this is curious, since in Darija negation is borne by the first marker *ma*, whereas the 2nd element (*ši, bu, ḥedd..*) bears the degree, the manner or the intensity: "not… *a single*, not… *at all*, not… *in the slightest*, not… *anyone*"; it is a quantifier and it never has a negative meaning:

> Concerning the origin of this marker, no conclusive explanation is available. However, two options may be envisaged: the first one relates to the Tuareg Berber verb *iba* (and variants) signifying meanings such as 'there is no', 'cessation of' and 'lack of', of which the derived form *ăba* is used in optative constructions. Texts from the Ayer region (Niger) prove the existence of this verb as a negation adverb, i.e. *ebəw* 'no, really!' The second option has to do with the Arabic nominal modifier and determinant *bu* (e.g. *bu ṛās* 'him with the head' > 'big headed one') that occurs in both Arabic and Berber data. (Lafkioui 2013: p. 79 n. 25)

Lafkioui adds: "Until now, the negation with *ma ___ bu* has only been attested, at least in a systematic way, in the region of Oujda (particularly in the city)" (Lafkioui 2013: p. 84), whereas the phenomenon had been described as early as 1983 and 1993 for Fez (Caubet 1983 and 1993).

The other question to be raised is: why go and look for a Tuareg or Nigerian Berber origin for marker *bu*, when the Arabic etymology seems obvious and self explicit?

6.8.3

In order to support this argument, I will examine ***the usages of particle bu*** – apart from negation – in M.A., which seem fairly straightforward, and well described, before going back to its developments in negation:

Bu as *locator of properties*, is used to attribute properties to a person or an object, to create nicknames based on defects (or more rarely, qualities), to create toponyms or names; it is built as a construct state (see Colin below), with *no determiner* on the second term: *bu kərš* [father-belly] "the big eater" or "the guy with a big belly".

Colin, in his dictionary[22] defines its roles as follows:

> Comme premier terme d'un état construit, bu sert à former de nombreux sobriquets, le plus souvent préjoratifs; le mot perd alors sa valeur primitive de père pour prendre le sens de: celui qui est caractérisé par…, l'homme à…, bu kərš: individu à gros ventre, le père gros ventre. Suivi d'un pluriel sans article dans les toponymes, bō-qrōn: à Rabat, bū-žlūd: à Fès.

Henry Mercier, in his dictionary of M.A., translates this usage of *bu* as "titulaire de…. l'homme à…" [lit. holder of…, the man with…], and he gives an impressive list of over a hundred expressions built on this construction (see Mercier 1951: p. 18). It includes names of illnesses or plants, euphemisms naming sexual body parts, proper names, nicknames etc. I will give examples taken from his list, trying to classify them:

- illnesses: *bu dḥas* "whitlow"; *bu fmim / bu friqiš / bu qmiqim* "foot and mouth disease"; *bu glib* "cholera"; *bu ḥellan* "hydrocephalus"; *bu ḥemrun* "meastles"; *bu idida* "one-handed man"; *bu kebbar* "adenitis"; *bu ṣeffir* "jaundice"; *bu sellum* "sciatica"; *bu šwika* "chickenpox, scarlet fever".
- plants: *bu denjal* "aubergine"; *bu ɛwida* "pear".
- animals: *bu ferṭiṭu* "butterfly"; *bu fesses* "gnat"; *bu mlis* "lizard"; *bu mqiṣa* "earwig".
- attributes (defects or qualities): *bu ḥdebba / bu ḥdibba* "hunchback"; *bu laḥya* "bearded"; *bu nif* "smart, clever"; *bu rejlin* "with big feet"; *bu šentuf* "with very thick hair"; *bu šlaġem* "man with a moustache"; *bu udina* "credulous"; *bu udnin* "with big ears"; *bu ujhayn* "hypocrite, two-faced".
- names (used as first names too): *bu jemɛa* "the Friday man, i.e. born on a Friday"; *bu šta* "the rain man".

Boujloud, "the guy wearing animal skins" on the day after the Eid al-Adha, is also the name of a gate in the medina of Fez, Bab Boujloud. In Casablanca and further South, it is called *Bu Bṭayen* (pl. of *bṭana* "untanned sheep skin with the wool"). Traditionally, in rural societies, on the day following the Eid a man would dress in the recently slaughtered sheep and rams' skins, paint his face black, and roam the place, whipping the air with a branch of leaves (for fertility and good luck): a very scary experience for many … Since 2010, there seems to be a revival of this tradition even in urban environements (see Caubet 2011 for Casablanca, and the numerous pictures posted by young people on Facebook at the time of the Eid in recent years …).

---

22. See Iraqui-Sinaceur (1993: vol. 1, p. 133), entry *bu/bo*.

This etymology of *bu* allows us to account for its usages ***in negation*** in several dialects. Since *bu* is used both in Berber and in M.A. (Lafkioui 2013), the contact phenomenon is undeniable, but how it developed is still the question. Lafkioui claims a borrowing from Berber into Arabic (see above[23]), but I would rather plead for ***the simpler hypothesis of the locator bu*** (*< father, see the gloss above), used both in the creation of nicknames and as a second element of the negation.

One has to account for two different constructions: one in Al Hoceima and Oujda, with an article, and another one in Fez, with *no article*. Lafkioui gives for Oujda Arabic a pattern "[*ma* + verbal predicative syntagm + *bu* + noun]", but, judging from her examples, it should rather be [*ma* + verbal predicative syntagm + *bu* + **article** + noun] (Lafkioui 2013: p. 80)?

As for the scope of the negation, in Fez, Taounil and Msek, it is an excluding negation (not a drop of, not an ounce of, etc.), whereas in Oujda it is described as "basic negation" (…) "associated with a determined subsequent object, regularly marked by the definite article in MAO. The negator *ma ___ bu* is used for basic negation in descriptive contexts (…). Therefore, its second element does not function as a tool that demarcates the negation scope and may occupy different positions in the syntagm." (…) "(78) ma šra-w bu l-ḥawli had l-ɛām 'They did not buy a sheep this year'", and "(79) ma ɛand-na bu ḍ-ḍəw f d-dāṛ l-qdīma. 'We do not have electricity at the old house'" (Lafkioui 2013: pp. 80–81).

These usages are quite different from those described for Fez, both for the syntactic construction with the article on the object, and for the scope (basic negation/strong excluding negation).

As for the El Hoceima data, they would need further enquiry; but as regards the etymology, since *bu* is found in very close contact with Berber (Taounil and Msek) as well as in much laxer contact (Fez and probably Oujda), and because of its other uses as a locator, I would definitely plead for the M.A. origin, and for the invariant behind the locator *bu*.

---

23. Lafkioui (2013: p. 53) draws a map showing how this *ma_bu* negation travelled from the Rif (including Al Hoceima) to Oujda, which does not take Fez into account.

6.9   Lexicon: Intrication and loans

6.9.1

The analysis of borrowings is more straightforward when it comes to words like *ibawen* "broad beans", which is also found for North Taza (see Colin 1921: p. 59), or *aqezzu*[24] "puppy", both borrowed directly from Berber.

6.9.2

It is more intricate when a word combines an Arabic stem and a Berber pattern, in new words like *tažəfnit* or *tasukkant,* combining Arabic *žəfna* "large tub, basin or dish", *sekkum* "wild asparagus", and the Berber feminine pattern *ta___t*. This pattern is very common in central Morocco, but limited to "abstract nouns of profession and personal characteristic"; Harrell (1962: p. 88) writes: "The abstract noun names either the profession or the abstracted personal quality associated with the meaning of the stem", which differs from the above construction; Harrell gives examples like "*tabennayt* '(profession, art of) masonry' or *tawekkalt* 'gluttony'"; I could add *tagnawit* "the art of Gnawa (rites and music)". On the contrary, the above examples are nouns designating concrete objects and not notions or concepts.

6.9.3

Finally I'd like to introduce a novel hybrid word which will bring us back to *bu* (see above); we discovered it in Msek. It involves *bu* and combines an Arabic pattern and a Berber lexeme; it was uttered by Y. (13) when Y. Thomas asked him about the varieties of almonds present in their orchard. He answered saying it depended on how hard their shells were and how one could break them,[25] defining three types:

> *waḥəd-n-nuɛ, kayəthərrəs ġir mən əl-fuʾ, wella mən l-iddayən; hada, ma-huwa kaythərrəs b-əl-idd, ma … dəġya kaythərrəs mən l-ḥjər; kayn waḥəd axor, ṣɛib baš thərrs-u; u kayn waḥəd axor, sahəl baš thərrs-u. hadak lli kaythərrəs deġya, kayqulu l-u* "**Bu-ġommāsi**".
>
> One kind, you can break just with your mouth, or with your hands … the other, it does not break with the hands, it … it breaks easily with a stone; there is another one, which is difficult to break; and there is yet another one, which is easy to break. That one, which is easily broken, they call it "***bu ġommāsi***".

---

24. Salem Chaker (personal communication September 2014) confirmed that "a*qezzu* est une formation expressive que l'on peut rencontrer sous des formes voisines diverses: *aqezzuḥ, aqezzun, aqjun*".

25. In reference to the work of Malou Delplanque (2011) in Bni Boufrah.

The noun *bu ġommasi* is formed with locator *bu*, plus a word formed with a combination of a Berber noun, *tuɣmas* "teeth", transformed through an Arabic intensive pattern (Measure II with the gemination of C_2: *ġommās*),[26] and a *nisba-i > bu ġommās-i* "the one which has the property of being breakable with the teeth". It was striking to note that in Bni Boufrah,[27] this kind of almond is named just *snan* "teeth", implying the same breaking process; a metonymy combined with a sort of euphemism or a semantic shortcut?

7.   Conclusion

The dialects we have briefly introduced here show how interesting it can be to work in cooperation with ethnobotanists; as for the enquiry methods, which is important for linguists involved in field research, we share the same approach of immersion. We linguists imposed the use of good recorders, because we were interested in the linguistic data and not just in the contents. We had to adapt to the use of translation, which is not easy when you need to record uninterrupted good sound quality. The ethnobotanists had to adapt to our spontaneous reactions to what was being said. This transdisciplinary approach has brought us novel data. The work on the borders between languages and their intrication is really worth pursuing for the younger generations that shared our fieldwork during the years 2013–2014 in the North West of Morocco.

As for the linguistic part, it is interesting to note once again that those marginal dialects that were predicted to disappear in the middle of the 20th century are still present. We tried to do a general overview of the Jbala region in 2012 with fifteen students and our volume (see Vicente et al. 2017) shows that the Jbala region has not evolved much in the space of a century (referring to Colin's study in 1921 in North Taza or Lévy-Provençal in 1922 in Ouargha, for example). Even with the development of electricity (and thus, television, which the Bni Itteft family has had since 2007), mobile phone and new technologies, it has not changed much in the last twenty years since Maghdad did her fieldwork under the supervision of Simon Lévy (1992).

---

26. Harrell (1962: p. 66) calls it "the nouns-adjective of profession and personal characteristic".

27. Malou Delplanque, personal communication.

List of abbreviations

ADJ	djective	IMP	imperative	PREP	preposi-tion
ADV	adverb	IPFV	imperfective		
CONJ	conjunction	LOC	locative	PRO	pronoun
DEF	definite article	M	masculine	PROX	proximal
DEM	demonstrative	N	noun	PRST	presentative
DIST	distal	NUM	numeral	PRVB	preverb
EXS	existential	OBJ	object	PTCP	participle
F	feminine	OBL	oblique	Q	question marker
FUT	future	PL	plural	REL	relative
HESIT	hesitation	PN	personal noun	SG	singular
IDP	independent	POSS	possessive		

References

Caubet, D. (1983). *La détermination en arabe marocain*, Collection ERA 642, Université Paris 7, Paris.

Caubet, D. (1993). *L'arabe marocain*, tomes I (Phonologie et morphosyntaxe) et II (Syntaxe et catégories grammaticales, textes), Paris: Peeters.

Caubet, D. (2002). Questionnaire de Dialectologie du Maghreb (d'après les travaux de W. Marçais, M. Cohen, G.S. Colin, J. Cantineau, D. Cohen, Ph. Marçais, etc.). In *EDNA, Estudios de dialectologia norteafricana y andalusi* 5: 73–92.

Caubet, D. (2011). De Facebook au Boujloud: chroniques de l'Aïd 2011 à Casablanca. éditorial du site du Centre Jacques Berque, novembre 2011: https://www.academia.edu/12408274/De_Facebook_au_Boujloud_chroniques_de_l_A%C3%AFd_2011_%C3%A0_Casablanca_

Caubet, D. (2016). The dialect of Msek – Beni Itteft (Al Hoceima), on the borders with Berber – revisited. In Grigore G. & Bițună G. (eds.), *Arabic Varieties: Far and Wide*. Proceedings of the 11th International Conference of AIDA – Bucharest (2015). Bucharest: Editura Universității din București. 163–173.

Caubet, D. (2017). L'importance des parlers du nord-ouest marocain dans l'histoire de l'arabisation du Nord de l'Afrique: revisiter des corpus recueillis dans la région en 1992–1995 sous la direction de Simon Lévy. In Á. Vicente, D. Caubet and A. Naciri (eds.), *La région du Nord-Ouest marocain: Parlers et pratiques sociales et culturelles*. Zaragoza: Universidad de Zaragoza.

Caubet, D. and Thomas, Y. (2017). "Céréales, fours, pains et levains dans la région d'El Hoceima: techniques et notes sur des parlers arabes à la frontière de la berbérophonie", in Y. Thomas, Y. Hmimsa, Á. Vicente and D. Caubet (eds.), *Jbala du Rif : Relations des sociétés Jbala à la nature, parlers et relation à autrui dans le Rif, Nord du Maroc*, Revue d'Ethnoécologie, Supplément 1, 2017. https://journals.openedition.org/ethnoecologie/3065

Cohen, D. et Caubet, D. (2000). Un questionnaire différentiel pour la dialectologie arabe. In *Oriente Moderno* XIX (LXXX), n. s., 1–2000, *Studi di Dialettologia Araba*. 1–23.

Colin, G. S. (1937). Les parlers: l'arabe. In *Initiation au Maroc, Institut des Hautes Etudes Marocaines*. 208–236.

Colin, G. S. (1921). Notes sur le parler arabe du nord de la région de Taza. In*BIFAO* 18. 33–119.

Delplanque, M. (2011). *Histoire évolutive de l'amandier en méditerranée. Regards croisé sur la domestication, dialogue entre la biologie et l'éthnobiologie*, Thèse de Doctorat, Université Montpellier II, sous la direction de Hélène Joly et Yildiz Thomas, Montpellier. 264 p.

El Alaoui, N. (2007). Une presse à huile au Maroc. In *Techniques & Culture* [En ligne], 48–49, mis en ligne le 20 juin 2010, consulté le 27 février 2016. URL: http://tc.revues.org/4541

El Jattari, K. (2017). Contact de parlers arabe et berbère en zone du Rif : Région de Bni Iṭṭeft. In Á. Vicente, D. Caubet and A. Naciri Azzouz (eds.), *La région du Nord-Ouest marocain: Parlers et pratiques sociales et culturelles*. Zaragoza: Universidad de Zaragoza. 315–326.

Harrell, R. S. (1962). *A Short Reference Grammar of Moroccan Arabic*, Georgetown University Press, Washington D.C.

Iraqui-Sinaceur, Z. (ed.). (1993). *Le Dictionnaire COLIN d'Arabe Dialectal Marocain*, Editions El Manahil, Rabat (8 volumes).

Kossman, M. (2017a). La place du parler des Senhaja de Sraïr dans la dialectologie berbère. In Á. Vicente, D. Caubet and A. Naciri Azzouz (eds.), *La région du Nord-Ouest marocain: Parlers et pratiques sociales et culturelles*. Zaragoza: Universidad de Zaragoza.

Kossman, M. (2017b). Berber-Arabic Contact. In *Oxford Research Encyclopedia of Linguistics*. Published online (accessed 21 July 2017) http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-232?rskey=hOxCDP& result=1

Lafkioui, M. (2013). Reinventing negation patterns in Moroccan Arabic. In M. Lafkioui (ed.), *African Arabic: Approaches to Dialectology*. Berlin: De Gruyter, Mouton. 51–93. https://doi.org/10.1515/9783110292343.51

Lévy, S. (1998). Problématique historique du processus d'arabisation au Maroc. Pour une histoire linguistique du Maroc. In P. Cressier, J. Aguadé, et A. Vicente (eds.), *Peuplement et arabisation au Maghreb occidental, Dialectologie et histoire*. Casa de Velazquez. 11–26.

Lévy-Provençal, E. (1922). *Textes arabes de l'Ouargha: Dialecte des Jbala, Maroc septentrional*. Paris: Ernest Leroux.

Maghdad, A. (1993). *El habla árabe en el Aduar de Msek (Textos, Traducciones, Notas y Compiendo)*, mémoire de licence d'espagnol non publié, sous la direction de Simon Lévy, Université Mohamed V, Rabat. 1992–1993.

Marçais, P. (1977). *Esquisse grammaticale de l'arabe maghrébin.* Paris: Adrien Maisonneuve.

Marçais, P. (1956). *Le parler arabe de Djidjelli.* Paris: Adrien Maisonneuve.

Marçais, W. and A. Guîga. (1925). *Textes arabes de Takroûna* (Textes, Transcription et Traduction annotée). Paris: Imprimerie Nationale.

Marçais, W. (1911). *Textes arabes de Tanger: transcription, traduction annotée, glossaire*. Paris: Ernest Leroux.

Marçais, W. (1902). *Le dialecte arabe parlé à Tlemcen – Grammaire, textes et glossaire*. Paris: Ernest Leroux.

Maurer, G. (1968). Les paysans du haut Rif central. *Revue de Géographie du Maroc, n. 14*. Rabat: Faculté des Lettres. 3–70.

Mercier, H. (1951). *Dictionnaire Arabe-Français*, Les Editions de la Porte, Rabat, 2ème édition .

Moscoso, F. (2003). *El dialecto árabe de Chauen (N. de Marruecos), Estudio lingüísistico y textos.* Universidad de Cádiz.

Natividad, E. (1998). Le dialecte de Chefchaouen. In P. Cressier, J. Aguadé and A. Vicente, *Peuplement et arabisation au Maghreb occidental, Dialectologie et histoire*, éds. Casa de Velazquez. 109–120.

Renisio, A. (1932). Étude sur les dialectes berbères des Beni Iznassen, du Rif et des Senhaja de Sraïr. *Grammaire, textes et lexique*. Paris: E. Leroux.

Thomas, Y. and Caubet, D. (2017). Savoirs paysans autour des huiles d'olive, (*zaytun*, *Olea europaea, var. europaea*) et d'oléastre, (*əl-bərri*, *Olea europaea, var sylvestris*) Rif, Nord du Maroc. *De la reproduction des arbres aux pratiques alimentaires*. In Y. Thomas, Y. Hmimsa, Á. Vicente and D. Caubet (eds.), *Jbala du Rif: Relations des sociétés Jbala à la nature, parlers et relation à autrui dans le Rif, Nord du Maroc*, Revue d'Ethnoécologie, Supplément 1, 2017. https://journals.openedition.org/ethnoecologie/3065

Thomas, Y. and Caubet, D. (2017). Usages des olives et techniques des huiles: des pratiques de cueillette à la table, à Ouezzane, Ain Mediouna et Msek au Maroc. In Y. Thomas, Y. Hmimsa, Á. Vicente and D. Caubet (eds.), *Jbala du Rif: Relations des sociétés Jbala à la nature, parlers et relation à autrui dans le Rif, Nord du Maroc*, Revue d'Ethnoécologie, Supplément 1, 2017. https://journals.openedition.org/ethnoecologie/3065

Vicente, Á. (2005). *Ceuta: une ville entre deux langues.* Paris: L'Harmattan.

Vicente, Á. (2000). *El Dialecto Árabe de Anjra, Estudion Lingüístico y textos*. Universidad de Zaragoza. 285 p.

Vicente, Á., D. Caubet and A. Naciri Azzouz (eds.). (2017). *La région du Nord-Ouest marocain: Parlers et pratiques sociales et culturelles*. Zaragoza: Universidad de Zaragoza.

Arabic on the Dahlak islands (Eritrea)

Marie-Claude Simeone-Senelle
CNRS, LLACAN-INALCO, Université Sorbonne Paris-Cité

This paper is devoted to the Arabic dialect spoken on the Dahlak archipelago of Eritrea, a variety of Arabic poorly documented so far. There are few studies on the Arabic varieties spoken on the African coast of the Red sea (Simeone-Senelle 2000b, 2002, 2005a–b, 2009; Kassim Mohamed 2012) but none of them has been dedicated particularly to Arabic as spoken on the islands. By revising previously published data about the Arabic variety spoken by islanders, I will attempt to assess the specific features of Dahlaki Arabic. After an overview of the archipelago and its sociolinguistic situation on the three inhabited islands, the main features of Arabic spoken on the islands will be compared with Arabic spoken as *lingua franca (ALF)* on the African coast of the Red Sea. The issue is to determine to what extend a distinction can be drawn between both Arabic varieties: Dahlaki Arabic and *ALF* of the coast.

Keywords: multilingualism, language contact, *lingua franca*, ʿAfar, Southern Semitic, Dahalik, Tigre, borrowing, codeswitching

0. Introduction

Before delving into the description of Arabic spoken on the archipelago, a general overview will be provided concerning its population and the sociolinguistic situation with some information about the Arab presence on the islands, the status of Arabic and its different varieties in contact. The second part of this article will be devoted to the description of phonology and phonetics, morphosyntactic and lexical characteristics of Dahlaki Arabic; ultimately, at the discourse level, some examples of codeswitching will illustrate the impact of contacts with two other languages, ʿAfar and Dahalik spoken on the islands.

All linguistic data have been collected during fieldworks in 1996 and between 2002 and 2006. They are unique but unfortunately not recent because since 2006 the scientific missions are no more authorized in this region. Moreover the linguistic situation is linked to the political one and it changes rapidly in a country where the

males serve for prolonged periods in the armed forces, far from their native region, and in contact with speakers of different mother tongues. Hence these data reflect a time-limited situation before 2007.

The corpus is based on spontaneous narratives and tales, their explanation and some translations of Dahalik in Arabic. They have been recorded with men, some women and young people, essentially Dahalik [dahālik] and/or ʿAfar native speakers. There are very few Arabic native speakers, all of them being bi- or multilingual. The corpus in strictly Arabic mother tongue is therefore too poor. In this context it is extremely hard to arrive at a clear delineation of features belonging to vernacular or vehicular Arabic within the variety spoken on the archipelago. Both Arabic varieties are entangled in one, named Dahlaki Arabic (= *DKA*). This name refers to the Arabic variety spoken on the islands, while Arabic lingua franca (= *ALF*) refers only to the vehicular varieties spoken along the Red sea coast.

1.   General overview of the archipelago

The Dahlak archipelago in Eritrea lies off the port of Massawa. It consists of more than two hundred islets and islands. Only four islands are permanently inhabited: the greatest one, Dahlak Kebir [dahlāk-kebīr] (= D.K.) lying in the centre of the archipelago, Nora [nōra], to the north of D.K. and Dehil [deḥil], to the north-west and close to the continent. The fourth one is Dese [deseʔ] to the south-west and very close to the Bori Peninsula (see Map 1). It will be not dealt with further because it is inhabited exclusively by native ʿAfar speakers. In 2006 the permanent population of the three islands added up to almost 3,000 islanders, with about 2,000 on D.K., 380 on Nora and 600 on Dehil.[1]

1.1   Sociolinguistic situation

An overview of the different Semitic and Cushitic languages in contact in the area, on the mainland, is given on Map 1. Three languages are in contact on the archipelago: Dahalik [dahālík], an Ethiosemitic language endemic to the islands, ʿAfar, a Cushitic language spoken in Eritrea, Ethiopia and Djibouti, and Arabic, mother tongue of a minority of islanders also used as a *lingua franca*. Dahalik is the predominant language with 1,680 native speakers among 3,000 islanders; ʿAfar

---

1.   The numeric data are not of official origin. They are estimations based on the speakers' information collected in 2004 and 2006 on the three islands. They were valid only for this period and certainly at that time there existed already some discrepancies between the results of the inquiry and the reality. We have no information about the population on the archipelago nowadays.
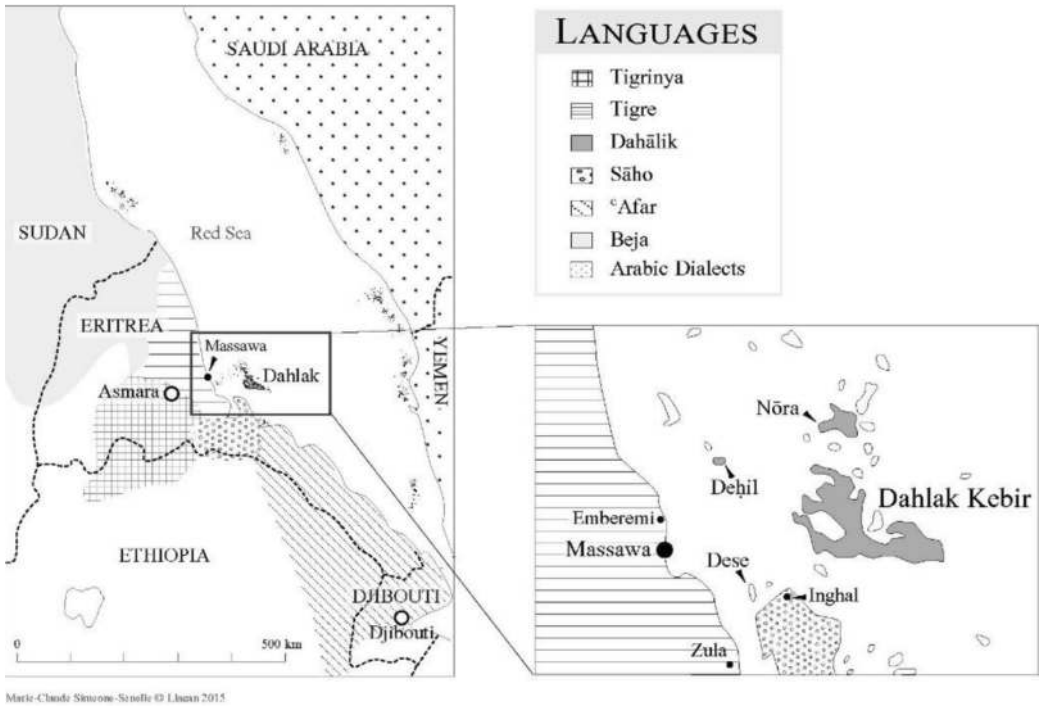
Map 1. Linguistic environment of the Dahlak Archipelago

is spoken by about 780 speakers, and Arabic is claimed as mother tongue by a minority of speakers: 350 on Dehil, 130 on Nora and 40 on D.K.

More specifically the languages are distributed as follows:

1. Dahalik is the majority language: 56.5% of islanders have it as mother tongue (60% on D.K., 60% on Nora, 42% on Dehil).
2. ʿAfar is the mother tongue of about 26% islanders (38% on D.K. and less than 5% on Nora).
3. Arabic speakers are a minority with 17.5% of islanders claiming to have Arabic as mother tongue. Nevertheless, the true situation is more complex because many islanders confuse the claimed identity (Arab) and the language spoken (Arabic). Many speakers of Arab origin say to be Arabic native speakers whereas they speak another language (Dahalik or ʿAfar) at home. This case is exemplified through Example (1). The man, recorded in a small village on D.K., was about 55 years old and was introduced as an Arabic native speaker.

(1) *ána   ʕárabi el=ʒins   ḥágg=i ʕárabi el=ʔásel   ʕárabi*
    PRO.IDP.1SG Arab   DEF=identity POSS=1SG Arab   DEF=origin Arab
    fī=l=bēt   natkéllem   dahlāk …
    in=DEF=house   1.IPFV.speak Dahlak(i)
    'I am Arab, my identity is Arab, the origin is Arab (…). At home, I speak Dahalik.'

Actually the number of islanders who claim to speak Arabic as their mother tongue or first language should be estimated to be about 2% on D.K., 35% on Nora and 58% on Dehil, where they represent a majority.

Everywhere, the prevalent situation is multilingualism. It was furthered by the Eritrean linguistic policy for promoting the education in all nine listed national languages and by favouring the contacts between all the national ethnic groups and their cultures. In the same district, the same village and the same family, people of different mother tongues are in contact. The following Example (2) is an illustration of this situation. The woman, an 'Afar native speaker of about 40 years old, answers in Arabic to the question in Arabic: 'What languages do you (SG) speak?'

(2) *nitkéllem   ʕárabi ḥábba ḥábba maʕa=l=ʕárabiye*
     1.IPFV.speak Arabic a_little a_little with-DEF=Arab_people
     maʕa=d=dáhlakiye   nitkéllem   dáhalik (sic)
     with=DEF=Dahlak_people 1.IPFV.speak Dahalik(for dahālik)
     u=maʕa=d=dénkaliye   nitkéllem   dénkali tigre wállah
     and=with=DEF=Dankali_people 1.IPFV.speak Dankali Tigre oh_God
     má=naʕref   kátir
     NEG=1.IPFV.know   much
     'I speak Arabic just a little, with the Arabs, with the Dahlak inhabitants, I speak Dahalik, and with the inhabitants of Dankaliya (coastal 'Afar region), I speak Dankali (i.e. 'Afar, in Arabic), as for Tigre (Ethiosemitic language spoken on the mainland), my Goodness, I don't know it much'.

With the exception of some very old women, nobody is monolingual on the islands. Any adult speaks at least two languages: her/his mother tongue plus a second language, and many people have a passive knowledge of a third or even a fourth. As school in the three islands is in Arabic, all educated children are acquainted with Standard Arabic.

The situation varies according to the island and the localization of the inhabitants of different languages within the same village. It also depends on the speakers' curriculum vitae, their sex, age, professional activities (trade, fishing) and their degree of schooling. In most villages, the linguistic communities interact on a daily basis. This situation of daily contacts is furthered by the short distances between the villages, and by the school gathering children from all communities.

Due to mixed marriages and shared daily life with children from other linguistic communities, many children actually have two first languages, and one of the two is always Dahalik. After the primary school some children go to Massawa to study (in Standard Arabic) and come back to their island for vacations. In Massawa, they are in contact with Tigrinya and, depending on the location of the school and the district where they are living during the scholar program, they have contact

with Tigre, ʿAfar and Saho (Cushitic). Obviously, this situation concerns much more the boys than the girls.

Concerning the women, they are in close contact with their neighbours in their village, but they are not used to travel outside the island, except for medical reasons or a rare visit to relatives in Massawa. With the exception of a few Tigre women married to islanders, none of them knows Tigre. Some Tigre women speak Tigre with their young children but all of them speak also ʿAfar or Dahalik or Arabic, depending on their husband's mother tongue and the neighbourhood. Like for men, the general situation depends on the villages. For instance on D.K. (Map 2a), in Dubʿullu, where the speakers of different languages are in daily contact within a small area, some ʿAfar women have a passive knowledge of the Dahalik language, to say the least. In other places many Dahalik women understand Arabic and ʿAfar, but they do not speak them fluently. I noticed that women who had different mother tongues (Dahalik and ʿAfar or Arabic) were able to talk together for hours, each speaking in her mother tongue but nevertheless understanding each other.

For men the situation is more complex. Because of their business and their contacts with the coastal mainland, many know Tigre and Tigrinya in addition to Dahalik, Arabic and ʿAfar. Through their professional activities some of them are also in contact with Saho, and understand a few English words. Among the oldest men a few know Amharic, the official language in Eritrea from 1959 to 1991, because they spoke it when they were inmates in the Derg government's prisons. Lastly they have some notions of Italian, acquired through their professional activities in Italian companies remaining in Massawa after1941, at the end of the colonial period. It is to be noted that on Dehil island, the nearest to the coastal mainland, men have very regular contacts with Emberemi (Map 1), the Tigre speaking area adjacent to their island. Besides Arabic and Dahalik, all of them therefore know Tigre and are able to speak it fluently.

The trade exchanges and the professional and personal relations with the eastern bank of the Red sea, an exclusively Arabic speaking area, are favouring the expansion of Arabic. On Dehil there are no ʿAfar and the Arabic and Dahalik native speakers live in two distinct but neighbouring villages (Dahret and ʿAd, Map 2b). The school, where children of both linguistic communities are grouped together, is half way between the villages.

1.2   A brief overview of Arabic presence on the Archipelago

Links between the African and Arabian coasts are attested since Antiquity and are without doubt much older. Concerning the Arabic language it was obviously spoken in the region before Islamization. The expansion of Islam in this area started from the
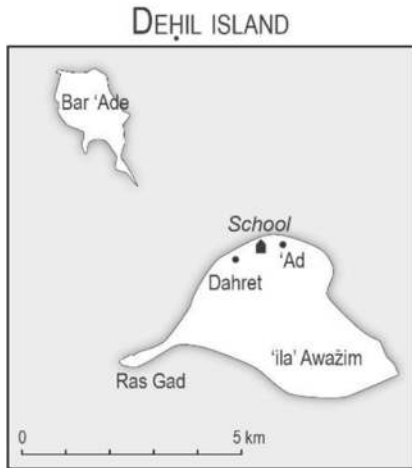
Map 2. (a) Dahlak Kebīr island; (b) Deḥil island; (c) Nōra island

Dahlak islands (Killion 1998: p. 265) and furthered the spread of Arabic. Nowadays, some islanders who claim to be Arabs because one of their male ancestors came from an Arab country are Dahalik native speakers (see above). The presence of Arabs on D.K. is attested since the beginning of Islam (Schneider 1983: p. 21), but it is recent on Nora where they state to have been here for three or four generations. Their ancestors migrated from different Arab countries: Egypt, Sudan, but mainly from Yemen (the Tihami coast, the Aden area, sometimes Hadramawt), and Saudi

Arabia (Farsan island). On Dehil, the Arabic native speakers claim to be Rashayda: indeed, in the nineteenth century the Rashayda tribe, including ʿAwâzim people, came from Saudi Arabia to settle along the coast north of Massawa. Evidence of their presence on Dehil island is provided by a Dahalik place like *ṣila-ṣawāzim* 'well of the ʿAwâzim people'.

In the state of Eritrea as a whole, only about 1% of the citizens are listed as having Arabic as their mother tongue (Abraha Wende, 2000, p.c.), their majority being Rashayda living on the northern coast, between Massawa and Sudan and very few of them, about 350, are settled on Dehil.

1.3   Status of Arabic

Modern Standard Arabic (= *MSA*) has an official status in Eritrea, where it is one of the three 'working languages' together with Tigrinya and English. As such it is used in written and oral forms in all official activities and in the media. *MSA* is taught in school as a second language in the country. The Eritrean educational policy stipulates that every mother tongue must be taught in the region where it is listed as vernacular. Because the Arabic varieties are not taken into account, it is *MSA* that is used as the medium of instruction and as a subject in primary schools on the northern coast, including the archipelago (see 1.4.1). Arabic is also the religious language on the archipelago, where all people are Muslims.[2] Religious, social and cultural usefulness confer prestige to Arabic, and usually the parents prefer for their children to attend the Arabic school, even in areas where it is not spoken. It is one of the reasons why speakers claim to be Arabs, wishing to be listed as Arabic native speakers. Indeed, for the speakers the Arabic language is a whole, and the prestige of Arabic furthers the use of their own Arabic variety.

1.4   Arabic varieties on the islands

1.4.1   *Modern Standard Arabic*
As the Eritrean educational policy was implemented while the Dahalik language was unknown, Dahalik is not listed among the national languages. On the Archipelago the policy has been aligned to the one applied to the northern coast, the Rashayda area, where Arabic is predominant, thus *MSA* is the only language of school education. Contacts with this Arabic variety through media is negligible because there are very few people listening to the radio or watching news on television in *MSA*.

---

2.   Muslims represent a little over 40% of the Eritrean citizens.

1.4.2   *Arabic dialects*

When islanders have a generator set and satellite dish, they invite their neighbours to look at TV programs: movies, varieties broadcast, television quiz shows from Arab countries (Egypt, Saudi Arabia, Yemen, Sudan, Emirates). For some children series of cartoons in Arabic are an effective means to learn the language.

Many people have relatives in Yemen and Saudi Arabia, the majority of men are fishermen or seafood traders and until 2008 they maintained regular relations with the other shore of the Red Sea, an Arabic monolingual area. During sea navigation Arabic is the Lingua Franca among the crew, which is mixed from the linguistic point of view.

ALF on the African coast results mainly from contacts with Arabic dialects spoken on the eastern bank of the Red sea, with Cushitic languages as ʿAfar and Saho, and with two Ethiosemitic languages, Tigre and Tigrinya spoken in Eritrea (Simeone-Senelle 2000b). By contrast, the range of contacts influencing *DKA* is limited to dialectal varieties of Arabic, including *ALF*, and to the mother tongues ʿAfar and Dahalik, the latter being dominant and endemic language on the islands.

The aim of this paper is to detect and assess the relevant characteristics of the *DKA* variety.

2.   Characteristics of Dahlaki Arabic

As expected, many characteristics are similar to the Arabic spoken as *ALF* along the African coast (Simeone-Senelle 2000b, 2002 and 2005b). Only some features are therefore examined among the most significant, at the level of phonology, phonetics, morphosyntax and vocabulary.

2.1   Phonology and phonetics

2.1.1   *Consonants*

The table of consonants is very close if not identical to that of *ALF* (Simeone-Senelle 2000b: pp. 157–164). Only a few features recorded on Dehil do differ and this can be explained by the predominance of Arabic native speakers and language contacts with Semitic languages (Dahalik, Arabic and Tigre).

Some phonemes occur only in special contexts, such as the literary code, and depend on the speaker, on the social status of his/her discussion partner, on the topic and literary value of his/her discourse. Moreover the same speaker can switch in the same text from one system to another, and use different realizations. The metadata must be taken into account insofar as the realization can vary with each idiolect.

Table 1.

	stop		nasal	fricative		emphatic		liquid		approximant
labial		b	m		f					
dento-alveolar	t	d	n	s	z	(ṭ)	(ṣ)	l	r	
palato-alveolar				ʃ	ʒ					
palatal										y
velar	k	g								w
uvular	q									
pharyngeal				ḥ	ʕ					
laryngeal				h						
glottal	ʔ									

2.1.1.1   *The emphatics*

As in the Arabic spoken on the coast, there are no voiced dento-alveolar emphatics. Only the two voiceless emphatics *ṭ* and *ṣ* are maintained in a few occurrences. They obviously have the same very unstable status as in *ALF* spoken on the coast.

The weak degree of pharyngealization may result now and then in the merging of *ṭ* into *t*, rarely into *tt* (Simeone-Senelle 2014: p. 72) and *ṣ* into *s* (for *ALF* see Simeone-Senelle 2005b: p. 269). On Dehil *ṭ* is clearly pharyngealized in some occurences. This articulation may be due to extended contact with the Arabic dialect mother tongue of the majority, and the influence of the Dahalik variety spoken on this island. Indeed in Dahalik of Dehil *ṭ* is generally retained as a pharyngealized phoneme, while on the two other islands, generally speaking, the emphatics are not maintained: *ḥeṭan* 'child, boy' on Dehil vs *ḥaʔan* on Nora and D.K, *ḥaṭa* 'marry' vs *ḥaʔa*. This process in Dahalik spoken on Dehil may itself be the result of contact with the Habab variety of Tigre, spoken on the coast, where *ṭ* may be pharyngealized (Elias 2005: p. 279).

In the same text in Arabic of D.K. (Simeone-Senelle 2014: pp. 70–71), recorded with a non Arabic native speaker who attended Arabic school on D.K., spoke Arabic very fluently and was in regular contact with the Yemenite coast, the word *maṭʕam* 'restaurant' has the same articulation in all the occurences of the word; the same with *rabaṭa=h* 'I tied it', *marbūṭ* 'hindered'.

/ṣ/ may be maintained as an emphatic in rare occurrences and in special contexts. It is emphatic in literary texts and usually with speakers in close contact with Arabic native speakers: *ṣaɣīr* 'small, young', is recorded on Dehil and in a tale (Simeone-Senelle 2014: p. 70), but *saɣīr* or *saqīr* elsewhere. It is to be noted that /ṣ/ is absent from ʕAfar, Dahalik, and the *ALF* of the African coast (Simeone-Senelle 2005b: p. 269). However, the toponym *mṣāwaʕ* 'Massawa' is always realized with an emphatic consonant as in the usual Arabic name *mṣawwaʕ* (the Dahalik name

is *mīwaʕ*). In addition, once in a while there may be a gradual contagion of the emphatic feature as in *mabṣūṭa* 'she is alright, well'.

2.1.1.2   *Interdentals*

As in the consonant systems of the three languages in contact, the interdentals are absent. When casually the voiceless one occurs, the word is repeated in the same text and 'corrected' with [t] or [s]: *keθīr* then *ketīr* and *kesīr* 'much'.

2.1.1.3   *Palato-alveolars*

The voiced fricatives /z/ and /ʒ/ together with the voiceless /ʃ/ are absent in ʿAfar, and unstable in Dahalik. In Arabic, native speakers of both languages often merge these consonants with the dental stop *d,* the velar stop *g* and the voiceless post-alveolar fricative *s*, respectively. When Dahalik native speakers realize fricatives, /ʒ/ is realized as a palatal [ɟ] and in some occurrences /ʃ/ as an alveo-palatal [ɕ]. This latter articulation is also attested for /s/ in some words in ʿAfar. Ultimately, some speakers, whatever their first language, use the variants in different occurrences of the same word.

   Example (3) is recorded with a single fluent speaker in Arabic whose mother tongues are ʿAfar and Dahalik:

(3) *ʃaɟara, ɕaɟara, saɟara*   and   *gedīra, ɟedīra, gezīra, ɟezīra*
     'tree'   and   'island'

As in *ALF* for the voiced consonant /ʒ/, the range of palatalization goes from [ɟ] to [y], as in many Arabic dialects in the area on both banks of the Red sea (Sudan, Yemen in Tihama and Hadramawt; cf. Simeone-Senelle 2000b: p. 161).

2.1.1.4   *Velars and uvulars*

In *DKA*, as in many Yemenite dialects and in Arabic dialects alongside the coast until Djibouti, there may be overlap between the voiceless fricative and plosive velars (*x ~ k*), between the voiced velar fricative and the uvular plosive (*ɣ ~ q* and *q ~ ɣ*). Finally, the uvular plosive is usually realized as a voiced velar (*q ~ g*; Simeone-Senelle 2000b: pp. 161–163).

   /q/ > [ɣ] on Dehil:

(4) *ɣarīb min=hum   nōra*
     close of=OBL.3PL Nora
     'Close to them, (there is) Nora island'

On D.K., in Jimhile:

(5)   *kīf   yeɣāsama=hu*
      how 3M.SG.IPFV.divide=OBJ.3M.SG
      'How does he divide it?'

However, in Durubbishet (D.K.), the same 'Afar and Dahalik native speaker[3] realizes [q] (6), [g] (7) or [k] (8):

(6)   *ḥikáya qábla   miya sána(t)*
      story   before 100 year
      'A story (that) may go back to 100 years'

(7)   *dágga=ni   bi=ʃuk   ḥágg=o*
      3M.SG.PFV.sting=OBJ.1SG with=prickle POSS=3M.SG
      'It stung me with its prickle (the ray)'

(8)   *ʒūz=ha   ma=yákrab   ʕánda=ha*
      husband=OBL.3F.SG NEG=3M.SG.IPFV.approach with=OBJ.3F.SG
      'Her husband does not approach her'

Only in specific contexts, usually literary ones, or with literate speakers, the etymological /ɣ/ and /x/ are retained and realized as velar fricatives. They say *ṣaɣīr* 'small', *xabbárt* <inform.PFV.2M.SG> 'you informed', *xalas* 'it's enough' (Simeone-Senelle 2014: pp. 69–70), beside *ṣ/saqīr, qāli* 'expensive', *kállast* 'I have finished', *lákam* 'shark'. In the same utterance, in different words, the speaker may have the three variants (Simeone-Senelle 2000b: p. 162). Unlike in *ALF*, in *DKA* the realization of */q/ as a uvular fricative [χ] or [ʁ] is rare.

The same Arabic native speaker on D.K., in Jimhile, says:

(9)   a.   *zēm(a) aχúl=la=k*
           As   1SG.IPFV.speak=to=OBL.2M.SG
           'As I tell you'

then, a little further:

(9)   b.   *hēze agúl=la=k*
           DEM   1SG.IPFV.speak=to=OBL.2M.SG
           'I tell you this'

In general the system is very fluctuating even with the same speaker. In (9b), to clarify his talk the speaker may switch to a more 'standard' variant of *ALF*.

---

3.   The 'Afar consonant system has not velar fricatives and no uvular; in Dahalik uvular fricatives are lacking and the distribution of the velar fricatives is strongly restricted.

2.1.2   *Vowels*

As in *ALF*, there are six short vowel realizations: *i, e, a, o, u, ə*. This latter is unstressed. With the exception of /ā/, the quantity is not relevant and it is linked to the stress and the consonant context. The systematic reduction of diphthongs to a long back or front vowel is maybe reinforced by contacts with Dahalik and 'Afar, both languages having no diphthongs.

(10)   *mūj* 'waves', *ēwa* 'yes', *bēn* 'in', *dē* 'like, as'

Some technical words are exceptions, such as *ḥaley* 'type of ray', *ʕáydān* (or *ʕídān*) 'pieces of wood'.

2.2   Morphosyntax

There is no morphosyntactic feature discriminating clearly the *DKA* from other varieties spoken in the area. However, some results concerning the verb (Simeone-Senelle 2000: p. 166; 2005: p. 271) should be mentioned.

2.2.1   *Verb*

2.2.1.1   *The IPFV conjugation*
Some speakers use in the IPFV conjugation the same form for both singular and plural first person: the personal index is *n-* and there is no number mark. This phenomenon is less rare than suggested in Simeone-Senelle (2000: p. 166; 2005: p. 271).

The Example (2) was recorded on D.K. in Melil with a female 'Afar native speaker, and the following (11) with a male Arabic native speaker. Only the man uses both forms for the 1.sg, in the same sentence.

(11)   *tigrinya   má=naʕref   má=darast=o*
        Tigrinya NEG=1SG.IPFV.know   NEG=1SG.PFV.learn=OBJ.3M.SG
        má=ʕaref
        NEG=1IPFV.know
        'I don't know Tigrinya, I didn't learn it, I don't know (it)'

2.2.1.2   *The PFV conjugation*
The 1rst and 2nd sg in PFV are identical and have a regular final *-a* (Simeone-Senelle 2005b: p. 271). This phenomenon, attested in some Yemenite dialects (Behnstedt 1985: p. 117) is more widespread in *DKA* than in *ALF*.

(12)   *samaʕta* 'I/you (sg) heard, understood'; *əntahēta* 'I/you (sg) finished';
        kúnta 'I was/you (sg) were'

Moreover, in the corpus collected on Dahlak Kebir there are some examples of this phenomenon for 3rd M.SG person, not attested in *ALF*:

(13)   *fi=baḥr artama*
       in=sea   3M.SG.PFV.throw_himself
       'He dived into sea'

(14)   *hāde insān   masaka   ed=dāra   bi=dahr=o*
       DEM person 3M.SG.PFV.hold DEF=house with=back=OBL.3M.SG
       'This man supported (the wall of) the house with his back'

2.2.1.3   *The system of TAM auxiliaries*
The system is more reduced than in *ALF*.

Only *kān* (in PFV) with verb in IPFV is used as aspectual and temporal marker with the same value (iterative, imperfect) as in *ALF*. The place of the auxiliary is variable before the verb (15), at the beginning of the sentence (16) or at the end (17):

(15)   *waḥad kān-idáwwur   abunuwās*
       one   AUX.3M.SG-3M.SG.IPFV.look_for Abu-Nuwas
       'Someone was looking for Abu Nuwas'

(16)   *kān   insān   yaʕni   maʕa   ḥimar   ḥaqq=o*
       AUX.3M.SG human indeed with   donkey POSS=3M.SG
       yemʃi
       3M.SG.IPFV.walk
       'In fact one human-being was walking with his donkey'

(17)   *ḥarb isáwwū   wu-igeṭṭəʕū   kānū*
       war 3PL.IPFV.do and-3PL.IPFV.slaughter AUX.PL
       'They were waging / they used to wage war and slaughtering people'

See below for *kān / ikūn* as a copula.

2.2.1.4   *TAM values*
As in *ALF*, concomitant or impending future can be expressed by a verb particle or a participle:

   ba= clitic to the verb in IPFV

(18)   *ed=dāra   ba=taṭiḥ   … ba=ʔaʒib   masāmir*
       DEF=house CONC=3F.SG.IPFV.fall … CONC=1SG.IPFV.bring nails
       'The house is about to fall down … I'll / I'm gonna bring nails'

The participle of some verbs of motion as (*raḥ*) or cognition (*ʕaraf*), has the same value:

(19)   *gál=la=ha*   *ana*   *rāyaḥ*   *mṣáwaʕ*
       3M.SG.say=to=OBL.3F.SG   PRO.IDP.1SG   PTCP.SG.go   Massawa
       'He told her 'I am going to Massawa''

Unlike in *ALF*, *bága* 'to stay' is not used as concomitance auxiliary. However the use of *bága* as enunciative particle 'and then', is exclusively attested in *DKA*. It marks the last element in an enumeration, such as a genealogical list or the final stage of a process (see examples in Simeone-Senelle 2000b: p. 272).

2.2.2   *Noun*
The morphological features are common to both Arabic varieties (*ALF* and *DKA*). The uncommon or unknown plural schemes are not specific to *DKA* insofar as they are attested on the islands and on the mainland coast. However plural forms similar to a dual are attested only in *DKA*:

(20)   a.   *elf*   *nafarēn*
            '1000   persons'
       b.   *áʒū*   *min* *bárra* *humma*
            3PL.PFV.come from   outside   PERS.PR.3PL
            'As for them, they came from elsewhere' (lit. they came from outside, they).

hum is more usual; here the use of *humma* reinforces the post-topicalisation of the subject.

2.3   Syntax

2.3.1   *The noun phrase*
The word order in noun phrase has no particularity: determiner/qualifier/modifier follows the determined/qualified/modified as in other Arabic varieties.

2.3.1.1   *The genitive construction*
The analytic construction is usual and not specific to *DKA*: both constituents are linked by a connective particle *ḥaqq*/*ḥagg* clitic to SUF.PERS.PR or DEF.noun.
       However another construction attested on the archipelago is not so common in Arabic: the determiner is a prepositional phrase with *men* and a suffixed personal pronoun. As in some Afrosemitic languages and in Soqotri (a Modern South Arabian language) the prepositional phrase is before the determined noun, not after it as usual in Arabic.

(21)   *mén=ni*   *awlād*
       from=OBL.1SG   children
       'My children/sons'

The synthetic construction is very rare, as in *ALF*. It is attested only with some names of kinship, part of the body or an element belonging exclusively to the referent of the determined noun.

(22) a. *áwled   awléd=hum*
        children children=OBL.3PL
        'Their children's children'
    b. *ísm=oh*
        name=OBL.3M.SG
        'His name'

When the speaker borrows vocabulary/phrases from another Arabic dialects. In (23) *ʒúwwat,* a feminine noun, may be borrowed from *guwwa* 'inside', preposition and adverb in Egyptian dialects (Hinds/Badawi 1986: p. 184).

(23) *ʒuwwat-el=baḥr*
        inside-DEF=sea
        'Into/within the sea'

2.3.2   *Nominal clause*
Different types of copula are used in the nominal clause.

2.3.2.1   *Copula of pronominal origin*
The 3rd person of the independent personal pronoun, inflected for gender only, is used as a predicative marker of the noun, with the value of existence and equative. The copula is usually at the end of the sentence; it does not occur in my *ALF* corpus, but inflected for gender and number it is attested in Dahalik.

(24) *hēdi naḥna   balád=na   hiya*
        DEM PRO.IDP.1PL country=OBL.1PL (PRO.IDP.3F.SG)COP
        'This one, for us, it is our country'.

The referent of *hiya* is *ʒezīra* (F.).

(25) *igūlu=l=u   safán igūlu=l=u   ḥaléy*
        3PL.IPFV.say=for=OBL.3M.SG *safan* 3PL.IPFV.say=for=OBL.3M.SG *ḥaley*
        nuʕwēn huwa
        types   (PRO.IDP.3M.SG)COP
        'They call it *safan*, they call it *haley*, they are different species (of rays)'

There are indeed many species of rays (for plural forms similar to dual see 2.2.2). Nevertheless the speaker's familiarity with different varieties of Arabic may explain the dual form as a codeswitching with standard Arabic.

A speaker from Dehil uses the suffix pronoun as copula:

(26) *ṣábər   kówwes ho*
      patience well   (PRO.IDP.3M.SG)COP
      'Patience, it is good'

2.3.2.2 *Prepositions used as copula*

As in other Arabic varieties in the Red Sea area, existence is expressed by *fī* = N., and possession by *ʕend*=SUF.PR referring to the subject of the predicate.

2.3.2.3 *Copula of verbal origin*

kān/ikūn is used to express existence, but without reinforcement by *fī* like in *ALF* (cf. Simeone-Senelle 2005b: p. 272).

(27) *ʕala=ḥásab sána ikún*
      on=account year (3M.SG.IPFV.be)COP
      'It depends on the year'

(28) *noxra   hēdi kēnet   ḥabs*
      Nokhra DEM (3F.SG.PFV.be)COP prison
      'This Nokhra (island) was a prison'

2.3.3   *Agreement*

As in *ALF*, the agreement in person, gender and number is very fluctuant, and often unpredictable, particularly with collective nouns. In *DKA* this feature is enhanced by the agreement systems in ʕAfar where the verb agrees in gender but not in number and remains in singular even with subject in plural. Moreover, in Dahalik the agreement with collective, and occasionally with plural, is irregular: either feminine or masculine, 3rd person plural or singular. Usually the counted noun is singular as in ʕAfar and in Dahalik. Only a few significant examples are noted below.

(29) Agreement of verb in number and not in person:

(29) *náḥna   yegūlu   kulla=nā   min   hīne*
      PRO.IDP.1PL 3PL.IPFV.say all=OBL.1PL from here
      'We say that all of us are from here'

(30) Agreement of the collective in 3M.SG or in 3PL, in the same sentence:

(30) a. *el=ḥarīm   yarqaṣ*
           DEF=women 3M.SG.IPFV.dance

and

(30) b. *el=ḥarīm*   *yarqaṣu*
       DEF=women 3PL.IPFV.dance
       'The women dance'

nās has usually plural agreement.

As in 'Afar, the ethnonym 'Afar (collective) has an agreement in the singular feminine:

(31) *ʕáfar*   *tugūl*
    'Afar_people 3F.SG.IPFV.say
    'Afar people say …'

(32) *el=ʕáfar*   *kulla=ha*
    DEF='Afar_people   all=OBL.3F.SG
    'All the 'Afar people'

2.3.4 *Sentence*

The constituent order seems attributable to the speakers' native languages (Simeone-Senelle 2005b: p. 272) and their familiarity with *ALF*. The instability in *DKA* may be reinforced by close contacts with 'Afar and Dahalik where the word order is usually SOV and Dependent clause before Matrix.

(33) *waḥad min el=baḥriya gāl*
    one   from DEF=seamen 3M.SG.PFV.say
    'One of the seamen spoke'

In a series of juxtaposed clauses, the order may vary from one clause to another:

(34) *lámma rēt=hum* ([rētuhum]) *ʃāfu-l=ʒuráḥ*   *el=kebda*
    when   1SG.PFV.see=OBJ.3PL   3PL.PFV.look-DEF=wound   DEF=liver
    ʃālu=ha   *u=ʒuráḥ*   *fākku=ha* …
    3PL.PFV.take=OBJ.3F.SG   and=wound 3PL.PFV.open=OBJ.3F.SG
    'When I saw them, they looked at the wound, they took the liver (of the fish) and they opened (my) wound' (to pour hot ray-fish liver oil into it).

In (35) the dependent is before the matrix, in both clauses the verb is at the beginning (V(S)O):

(35) *namma xāfu*   *min ed=dargi nāzahu*
    when   3PL.PFV.fear from DEF=DERG 3PL.PFV.keep_away
    min=ha ([minnaha]) *ilē nōra*
    from=OBL.3F.SG   to   Nora
    'When they were afraid of the DERG (the Ethiopian military junta), they departed from it (Nokhra island) to Nora (island)'

3. Vocabulary and borrowings

3.1 Vocabulary

3.1.1 *Ideophones*
As in 'Afar and Dahalik, there are many occurrences of ideophones but unlike in these languages they are not integrated into a verb phrase. They are an explanatory, iconic marker of the predicate.

(36) a. *tubtubtub yermu* 'plop, plop, plop they throw (in sea)'
 b. *tʃep fi-baḥr artama* 'splash! he dived (splash, in the sea he dived)'

3.1.2 *Interferences with Arabic dialects of the area*
There are numerous lexical interferences with Arabic dialects of the Red sea area with an obviously significant influence of the Yemenite dialects:
 They use *əʃti* (in the IPFV) to mean: 'to want'

(37) *eʃ tíʃti* 'What do you want?'

or in a broader meaning including possibility and eventuality:

(38) *fī=hōri   kaman tíʃti   miya-wu-xamsīn kílo*
 EXS=hori also   2SG.IPFV.want 100-and-50   kilo
 'There are also hori (a type of ships), you can load 150 kilos'

bɣa 'to want', usual in Hadramawt is used occasionnaly as in Tihama (*cf.* Behnstedt 1985: p. 202). However, the interrogative *eʃ* is not used on the western coast of Yemen.

(39) *eʃ   tábɣa*
 what.Q 2M.SG.IPFV.want
 'What do you want?'

'To come' is expressed by *agā / aʒā*, but *ata* usual in *ALF* is not attested in the corpus.
 rāʔe 'see', is used as in the Tihama of Yemen (Behnstedt 1985: pp. 194–199), in the *DKA* data, it is always in the PFV besides *ʃāf* 'see and look', in the PFV and IPFV (cf. 34).

(40) *lámma rēt=hum*([rētuhum])   *ʃāfu-l=ʒuráḥ*
 when   1SG.PFV.see=OBL.3PL   3PL.PFV.look-DEF=wound
 'When I saw them, they looked at the wound'

For *mōya, mō* 'water', *delḥin, deḥḥin* 'now', and for *bēn* 'inside', as in some Yemenite dialects (Jastrow 1980: p. 127), see Simeone-Senelle (2000b: p. 176).

For 'when', besides *lámma, námma* with nasalisation by some speakers (as on the coast) [lammã], [nammã], some use *émma* and *man.*

3.2   Lexical borrowings

For centuries the islanders have been in closer contact with the Arabian coast and generally speaking they are more focused on the eastern bank of the Red sea than on the mainland coast of Africa. Their familial relations and professional activities are directed essentially to the Yemen coast where people are all native Arabic speakers. So, the halieutic and fishery vocabulary, fish names, fishing techniques, boat pieces, boat names, has been borrowed mainly from Yemenite dialects. Dahlak Archipelago has been an important trade centre in the Red sea and many lexemes belonging to different semantic fields can be detected as etymologically loaned not only from Arabic dialects of the Red sea area (Egypt, Sudan, Saudi Arabia, Yemen) but from many other foreign languages. The lexicon reflects the past and the present history of the archipelago and the region: traditional and trading activities, movement of populations, colonisation. They are loanwords (Manfredi et al. 2015: p. 284) and as such they are totally integrated into the lexicon. They are common to the three languages spoken on the archipelago (Dahalik, 'Afar and Arabic) and along the African coast of the Red sea. For many of them, despite a high degree of morphophonological integration in the recipient language, the etymology is clear. Besides Arabic dialects, there are primarily two donor languages: Italian and English. Unlike in Dahalik, there are no obvious loanwords from Ethiosemitic (Tigre, Tigrinya or Amharic).

(41)   From Italian:
    bonsola (*bussola*) 'compass';
    basta (*pasta*) 'any pasta';
    bāni (*pane*) 'white bread';
    kōkiyan (probably *conchiglia* 'seashell') 'mother-of-pearl, any seashell that produces mother-of-pearl'.

(42)   From English:
    dūʃ (*dish*) 'satellite dish'
    itrik (*electric*) 'flashlight'
    fībər (from *fibre(glass)*) 'fibreglass boat'
    tan (*ton*) 'ton'
    kinīn (*quinine*) 'any pill'

kinīn could have been borrowed in the area through Arabic (same form, same meaning).

tan may be borrowed from English or from Arabic (**ṭun/ṭon*).

fībər: on the archipelago, in *DKA* the diphthongs are reduced, *ay > ī*, so the word is here probably borrowed from English through an Arabic dialect, as in Yemen where *faybər* and *fībər* are attested. For this lexical item see Manfredi et al. (2015: p. 289).

Although it is an integral part of the lexicon, it is worth focusing on the item *ʃamaṭri* 'cinnamon'. As far as I know it is attested only in *DKA*. Formally it is a relational adjective (*nisba*), literally meaning 'from Sumatra'. The form with *ʃ* could reflect a very old borrowing, if we consider the transliteration of the island's name as *shamuthera* by Niccolò de' Conti, a Venetian merchant in the 15th century (Yule and Burnell 1986: p. 866). As for *ṭ* attested in the Arabic name *sūmaṭra*, it should be equivalent etymologically to an aspirated stop transliterated *th*. However, the donor language of the loan is unknown and the origin of the word remains unclear. Nevertheless this name provides valuable information on the particular variety of cinnamon which has been imported on the Archipelago. It should be not *Cinnamomum verum* or *Cinnamomum zeylanicum* originating from Ceylon/Skri Lanka (Al-Munjid 1975: p. 622), the name of which is based on the root *<q/krf>* in Semitic and Cushitic languages of the area: Arabic *qirfa*, Tigrinya *qafra*, ʿAfar *kafrá,* Somali *qorf.* In his comments (in Ibn al-Bayṭar, t.II, 70), Leclerc explains that "le cinnamome compte plusieurs espèces, désignées par les noms des pays où on le trouve". *ʃumāṭri* is probably another species such as *Cassia cinnamom* or *Cinnamom aromaticum,* originating from the Indonesian islands (Moluccas and Sumatra) and from China, as evidenced by the name *dār ṣīnī* 'cinnamome' in Arabic, borrowed from Persian (lit. 'wood or bark from China', Renaud and Colin 1934: p. 51). Moreover the name suggests that the transit through Dahlak or the import on the archipelago of the cinnamon from Sumatra goes back to ancient times,[4] well before the arrival of the Ceylon or Skri Lanka variety on the market.

4.   Codeswitching

All the male speakers are multilingual and they use codeswitching towards Dahalik and ʿAfar. In narrative texts there are no examples of switching towards Tigre. Except for the use of some 'dual' forms (Examples (20a–b), (26)), the switching to another Arabic variety has not been considered here. Generally speaking, speakers switch more frequently to Dahalik than ʿAfar.

---

4.   Renaud and Colin (1934: p. 130): "… les Anciens n'ont connu que l'écorce du *Cinnamomum aromaticum* (…) qui fournit la cannelle de Chine. L'Arabie en était le marché dans l'Antiquité."

As one speaker on D.K. emphasized:

(43) *lámman agúl=l=ik*   *qesmēn qésəm dáhlaki*
when   1SG.IPFV.speak=to=OBL.2F.SG part.DU part   Dahlaki
wu qésəm ʕarábi
and part   Arabic
'When I speak to you, there are two parts: one part in Dahlaki (= Dahālik
language, in Arabic), one part in Arabic'

4.1 Towards Dahālik

In (44), the speaker, after a brief pause, uses the Dahalik number *kile* to explicit
DKA nafāren, ambiguous form (dual and plural of *nafar*). In *DKA* and 'Afar the
counted noun remains in SG.

(44) *nafāren, <u>kile</u> náfar*
persons two person
'People, two persons'

In (45) CSW appears when the Arabic-'Afar native speaker stumbles over the Arabic
word (*waladat*), then he starts again but in Dahalik.

(45) *ḥurmát=i   wal…   <u>wadaʕat</u>*
wife=OBL.1SG ga(ve) 3F.SG.PFV.give_birth
'My wife ha …has given birth'

Here CSW is favored by possible interference with *waḍaʕat* < wḍʕ > in Arabic (same
meaning as *waladat*) while *wadaʕat* only is used in Dahalik to mean 'give birth'.
Two types of CSW are attested.

4.1.1 *Intra-sentential CSW*
Usually inside the simple sentence the CSW concerns the predicate (cf. above
Example (45).

(46) *(bir) tult-miya-u-settin   <u>elle – dib=o</u>*
(well) three-hundred-and-sixty 3M.SG.PFV.have – in=OBL.3M.SG
'There are 360 wells in this place' (Lit. It (the place) has 360 in it)

In Dahalik {*elle* (Verb) – *dib* (PREP.in)=SUF.PR} expresses existence or possession.

4.1.2 *Inter-clausal CSW*
In complex sentence (47), the CSW concerns the verb of the matrix at the end of
the sentence, after the reported speech, as in Dahalik and 'Afar.

(47) *ʔármi*   *ʕáfeš*   *ballu*
      throw.IMP.SG luggage   3M.SG.PFV.say.OBL.3M.SG
      'He told him: 'Throw luggage!''

The hybrid form *ballu* may result from overlapping Dahalik and Arabic in quick talk: *ballu* </bilā=hu/ in Dahalik, for *gállu*</gāl=l=u/ in Arabic, 'he told him'.

4.2   CSW towards ʕAfar

Only intra-sentential CSW are attested in the corpus

(48) *áh*   *wáḥad* *ʕáfari*
      DEICT=ASS one   ʕAfar
      'This one is 'Afar' / 'This one it is an 'Afar'

In ʕAfar the existential copula *áh* (DEICT=ASS predicative marker) is the first constituent of a noun clause.

(49) *taɣribān kébbi*   *kam ʕam yekkee*
      about   before.OBL.1SG how   year 3M.SG.PFV.happen
      'It happened approximately how many years ago before my birth?'

As in ʕAfar, the (ʕAfar) verb is at the end of the sentence and in the interrogative form (lengthening of the final vowel).

5.   Conclusion

On the archipelago the situation results irrevocably in merging and levelling of both varieties of Arabic, vernacular and vehicular, into the Dahlaki Arabic variety. This latter is not, strictly speaking, deeply different from the variety spoken on the coast and both have many common features. However some specific features are detected in the corpus recorded ten years ago.

The influence of Dahalik is not as significant as it could be expected. The few features spotted as being from Dahalik language are actually just remnants tending to disappear, levelled by the increasing influence of Arabic (*ALF* and other varieties) and the very endangerment of Dahalik. As for ʕAfar, it has neither more nor less impact on this Arabic variety than on *ALF* of the coast of the Red Sea.

Acknowledgment

List of abbreviations

ALF	Arabic Lingua Franca	IPFV	imperfective
ASS	assertive	INT	interrogation
AUX	auxiliary	M	masculine
BOR	borrowing	MSA	Modern Standard Arabic
CON	connective	NEG	negation
CONC	concomitant	OBJ	object
COP	copula	OBL	oblique
CSW	codeswitching	PFV	perfective
DEM	demonstrative	PL	plural
DEF	definite article	PRO	pronoun
D.K.	Dahlak Kebir	PREP	preposition
DKA	Dahlaki Arabic	PTCP	participle
DEICT	deictic	Q	question
DU	dual	SG	singular
EXS	existential	~	alternation
F	feminine	=	clitic boundary
IDP	independent		

References

Al-Munjid fī al-lugha المنجد في اللغة (1975). Beyrut: Dār al-Mashreq.

Behnstedt, P. (1985). *Die nordjemenitischen Dialekte. Teil 1: Atlas*. Wiesbaden: Dr. Ludwig Reichert.

Elias, D. L. (2005). *Tigre of Habab: Short Grammar and Texts from the Rigbat People*. Thesis. Department of Near Eastern Languages and Civilizations. Harvard University: Cambridge, Massachussetts.

Ferguson, C. (1970). The role of Arabic in Ethiopia: A sociolinguistic perspective. In J. E. Alatis (ed.), *Bilingualism and Language Contact*. GURT. (Georgetown University Round Table). University School of Languages and Linguistics: Georgetown. 355–372.

Hinds, M. and E. Badawi. (1986). *A Dictionary of Egyptian Arabic. Arabic-English*. Beirut: Librairie du Liban.

Ibn al-Bayṭar, A. A. (1881). *Traité des simples* (L. Leclerc Trad and Comment), t. II. Paris: Imprimerie nationale.

Jastrow, O. (1980). Text aus Yarim, Nordjemen & Texte aus Giblah (Jiblih), Nordjemen. In W. Fischer and O. Jastrow (eds.), *Handbuch der arabischen Dialekte*. Wiesbaden: Harrassowitz.

Kassim Mohamed, S. (2012). *Le parler hakmi de Djibouti, arabe vernaculaire de la capitale*. Thèse d'université. Inalco-Paris.

Killion, T. (1998). *Historical Dictionary of Eritrea* [African Historical Dictionaries, No. 75]. Lanham, Md., & London: The Scarecrow Press.

Manfredi, S., M.-C. Simeone-Senelle, and M. Tosco. (2015). Language contact, borrowing and codeswitching. In A. Mettouchi, M. Vanhove, and D. Caubet (eds.), *The CorpAfroAs corpus of spoken AfroAsiatic languages. Corpus-based Studies of Lesser-described Languages.* Amsterdam-Philadelphia: John Benjamins. 283–308.

Renaud, H. P. J. and G. S. Colin. (1934). *Tuḥfat al-Aḥbāb. Glossaire de la matière médicale marocaine*. Paris: Paul Geuthner.

Schneider, M. (1983). *Stèles funéraires musulmanes des îles Dahlak (Mer Rouge)*. I. *Introduction, Documents et Indices*. Le Caire: IFAO.

Simeone-Senelle, M.-C. (2000a). La situation linguistique dans le sud de l'Erythrée. In H. E. Wolff and O. D. Gensler (eds.), *Proceedings of the 2nd World congress of African Linguistics, Leipzig 1997*. Cologne: Rüdiger Köppe. 261–276.

Simeone-Senelle, M.-C. (2000b). L'arabe véhiculaire parlé en Erythrée sur la côte de la mer Rouge, de Massawa à Rahayta. In L. Bettini (ed.), *Oriente Moderno*, n.s. XIX (LXXX), 1: 153–181.

Simeone-Senelle, M.-C. (2002). L'arabe, langue maternelle de citoyens djiboutiens du nord de la République de Djibouti. In A. Youssi, F. Benjelloun et al. (eds.), *Proceedings of the 4th Conference of the International Arab Dialectology Association (AIDA), Marrakesh, Apr. 1–4. 2000. In Honour of Professor David Cohen* Rabat: AMAPATRIL. 140–150.

Simeone-Senelle, M.-C. (2005a). Djibouti / Eritrea. In K. Versteegh et al. (eds.), *Encyclopedia of Arabic Language and Linguistics* vol.1, Leiden: Brill. 654–659.

Simeone-Senelle, M.-C. (2005b). The Horn of Africa. In K. Versteegh et al. (eds.), *Encyclopedia of Arabic Language and Linguistics*, vol. 2, Leiden: Brill. 268–274.

Simeone-Senelle, M.-C. (2006). Some characteristics of Dahalik, a newly discovered Afro-Semitic language spoken in Eritrea. In S. Uhlig. et al. (eds.), Proceedings of the 15th International Conference of Ethiopian Studies. Hamburg, July 2003. Wiesbaden: Harrasssowitz. 861–869.

Simeone-Senelle, M.-C. (2008). The specificity of the Dahalik language within the Afro-Semitic languages [G. Lusini (ed.) *History and language of the Tigre-speaking peoples*]. Studi Africanistici. Serie Etiopica 7. 127–145.

Simeone-Senelle, M.-C. (2009). The Linguistic Situation on the Dahlak Islands in Eritrea [Proceedings of Red Sea Project IV. Held at the University of Southampton. September 2008]. *Connected Hinterlands*. BAR International Series 2052. London: Archaeopress. 61–68.

Simeone-Senelle, M.-C. (2014). Le chat borgne. Un conte en arabe véhiculaire des îles Dahlak (Erythrée). In L. Casini, P. LaSpisa, and A. R. Suriano (eds.), [The language(s) of Arabic Literature. Un Omagio a Lidia Bettini]. *Quaderni di Studi Arabi N.S.* 9. Rome: Istituto per l'Oriente Carlo Alfonso Nallino. 67–80.

Yule, H. and A. C. Burnell (1986 [1886]). *Hobson-Jobson. A Glossary of colloquial Anglo-India Words and Phrases*. London: Routledge and Kegan.

Ḥassāniyya Arabic in contact with Berber

The case of quadriliteral verbs

Catherine Taine-Cheikh

Lacito (CNRS, Universités Paris III-Sorbonne nouvelle and Inalco)

Generally speaking, it is at the edges of the Arabic speaking world that one finds the most borrowings and where the influence of contact on the internal development of Arabic is most visible. Although Mauritanian Ḥassāniyya is an exception to this general trend (Taine-Cheikh 1994, 2007), the dialect has nonetheless retained traces of the region's past and namely of the very gradual disappearance of Zenaga Berber.

My goal here is to assess, based on the study of a particular lexical sub-category (verb forms with quadriliteral roots), the influence Berber may have had on lexical formation in Ḥassāniyya Arabic.

1. Introduction

1.1 Ḥassāniyya Arabic

Ḥassāniyya is the Arabic dialect spoken in Mauritania and more broadly in West Africa, in the (Sahelo-)Saharan area – from Guelmin (Morocco) and Tindouf (Algeria) to the Senegal river, and from the Atlantic ocean to Timbuktu (Mali) or even East Niger. This is a region where, even after the gradual arrival of Arabic speaking tribes from the North, the inhabitants of the Sahara long continued to speak Berber. The original Berber dialects spoken in the area were distinct from the Tuareg languages.[1]

In Mauritania, the survival of Zenaga Berber is threatened by the continuous expansion of Ḥassāniyya. The number of speakers has steadily diminished over

---

[1] Zenaga was long considered the only representative of the South-West group, this view has changed with the study of Tetserret Berber (in West Niger) and certain varieties of Songhai (Mali, southern Algeria).

the past four centuries to reach a few thousand, all bilingual, in the Gebla region (in the South-West).

Today there are more than four million Hassanophone speakers (of which two-thirds are to be found in Mauritania). The dialect's origins are attributed to the Bani Maᶜqil (late 13th-early 14th century); it shows remarkable unity despite the size of the territory where it is spoken. Its peculiarities place it among the Maghreb and/or so-called 'Bedouin' dialects. It also has many specificities in all domains, i.e. phonetics, phonology, morphosyntax, the lexicon and semantics.

Scholars of Arabic often ascribe these peculiarities to contact with Berber and Sub-Saharan African languages. While 'Africanisms' in Ḥassāniyya (and in Zenaga) are rare (and only to be found in the lexicon),[2] the influence of Berber is obvious but not easy to measure.

1.2   Reciprocal influence between Berber and Arabic

Several centuries of contact between Berber and Arabic has had a deep impact on Berber, in particular on northern dialects (Kossmann 2013).

Inversely, the Arabic dialects of the Maghreb have been influenced by the Berber substrate, although it is not always easy to determine which forms are borrowings (see e.g. the case of the preverbal form *lā* in Chefchaouen Arabic, see El Aissati 2006: pp. 294–5).

I examine elsewhere the interferences between Ḥassāniyya and Zenaga from a morphosyntactic perspective and conclude that convergence, less widespread than expected, often takes on the form of parallel developments (Taine-Cheikh 2008a).

At the lexical level, it is difficult to establish a percentage for the entire Ḥassāniyya vocabulary (Taine-Cheikh 1988–1998). Of course a word's Arabic origins are obvious when there is formal and semantic convergence with Classical Arabic and/or a set of dialects of diverse geographical origins (not only Northern African). Similarly, Berber origins are quite certain when a non-Arabic lexeme (in accordance with the above definition) is found in several Berber dialects (or languages). Beyond these two cases, there are various intermediate possibilities where the original language is uncertain, for example if the lexeme is only found in Zenaga Berber or if, in Arabic, it is only found in Northern Africa, because in both cases the source language could be either Arabic or Berber.

Luckily there are also phonological and morphological clues to inform conclusions.

---

2.   My remark here bears on linguistics. Of course there is much to be said on e.g. music and the social institution of griots.

A. Borrowings in Arabic from Berber often have the following characteristics:

i.   presence of the emphatic consonant /ẓ/ to which, in the case of Ḥassāniyya, one may add the palatalized consonants /dʸ/, /tʸ/, and perhaps even /nʸ/;[3]
ii. a specific syllable structure allowing for the presence of vowels in open syllables, and noun patterns rich in vowels, with a tendency to pronounce vowels as short, mid or long depending on where they are in the word (and independently of their length in Zenaga, see Taine-Cheikh 1997);[4]
iii. noun affixes comparable to those found in Berber, namely
     –   vowel prefixes, usually *a-/ā-* in the singular and *i-/ī-* in the plural, although there are also others, cf. *āršân* PL *īršîwən* 'sump'; *īggîw* PL *īggâwən* 'musician (griot)', *ādlägân* (PL) 'beans';
     –   a prefixed marker in *t-* for feminine nouns, with a suffixed *-t* often being added (especially in the singular), ex. *tādît* PL *tīdâtən* 'container for milking';
     –   a plural marker in *-n* for nouns (often M.PL *-ən* vs. F.PL *-än*);
iv. regular omission of the definite article before borrowed nouns, whence *īggîw* 'a/the griot'.

B. Borrowings in Zenaga from Arabic may have the following characteristics:

i.   Arabic nouns are often borrowed along with the definite article *al-* (the *l-* being assimilated to the first noun root consonant in some cases – so-called solar letters because they follow the model of *äš-šäms* 'the sun', *ž* included);
ii. borrowed verbs often show a specific pattern in *yaCa(Ca …)Ca(h)* and have highly reduced vowel alternation.[5]

All of these characteristics will be highly useful for this research, even though none of the criteria are fully decisive. The Berberized form *ägārâž* of the Ḥassāniyya borrowing of French *garage* is a telling exception to the rule. Let us further note that in Zenaga, a Ḥassanized form can coexist with another form, whether verbal or not, which is not Ḥassanized (e.g. *yäššäyväräh* variant of *yäššäffär* 'be the neighbor of'; *yängārä* 'leave in the opposite direction' of which the masdar is *əngīri*).

---

3.   These phonemes are found in neighboring Sub-Saharan African languages, but their presence in Zenaga is undeniable (see e.g. the passive with the prefix *Tʸ*).

4.   Long vowels are stressed (noted by a circumflex accent).

5.   The imperfective is formed simply by changing the prefix vowel *yəCa(Ca …)Ca(h)* and has no specific pattern for negative conjugations.

Lastly, we will see that while the Berber factitive marker is often recognizable, despite the multiple forms the morpheme 's' can take in Berber (and namely in Zenaga),[6] it can also be affixed to stems of Arabic, and not only Berber, origins.

In order to determine which quadriliteral verbs could have Berber origins, I will take into account all of these criteria.

Some remarks on the specificities of the dialects considered:

i.   In Ḥassāniyya, /f/ is generally pronounced voiced [v]. This Bedouin dialect has retained interdental consonants, and the *qâf* is pronounced voiced *g*, but some speakers pronounce the *ɣayn* as *q*.

ii. Zenaga is characterized namely by two specific (among Berber) regular changes: /l/ > /y/ and /ɣ/ > /ʾ/. It is a 'spirantizing' language where stops are realized as fricatives in intervocalic position.[7]

1.3   Quadriliteral verbs

Languages have a strong tendency to borrow isolated lexemes, in particular when they are nouns referring to *reales*. Verbs are less frequently borrowed, which has given rise to debate on the specific structural properties of borrowed verbs. So, Wichmann & Wohlgemuth (2008) assert: "we distinguish for major types – the light verb strategy, indirect insertion, direct insertion, and paradigm transfer. […] We count as loan verbs all items that function as verbs in the source language, even if they are treated as nouns (and are subsequently verbalized) in the borrowing language (a common phenomenon). In contrast, we exclude case where the item functions as a noun in the source language […]".

I will return below to the question of the syntactic category of borrowings in the source language, however my study, bearing on verbs in Ḥassāniyya (or at least those with quadriliteral roots), examines their lexical rather than their morpho-syntactic characteristics.[8]

---

6.   The most common pronunciations in Berber are alveolar (*s, z, ṣ, ẓ*) or postalveolar (*š, ž*) – whether the consonant is a geminate or not. Zenaga shows further complication in that there are frequent differences in pronunciation across simple and geminate consonants (Taine-Cheikh 2008b).

7.   The notations *ⱬ, ⱬ̣* and *ẕ̌* denote spirant and/or approximant pronunciations of the consonants *z, ẓ* and *ž*, specific to Zenaga. The *ġ* notes here the 'soft' pronunciation of the *g* in Tuareg.

8.   However, the "light verb construction" (nominal + light verb), rare in Arabic, has been noted in a few dialects. For example in Sason Arabic, through contact with Kurdish and Turkish (Akkuş 2016: p. 39). In Ḥassāniyya, it appears to me to be little used, but one of my consultants refused *kälkäl* in favor of *dâr tākəlkâlət (əl vlân)* for 'apply a poultice (on someone)'.

Functionally speaking, all quadriliteral verbs in Ḥassāniyya follow the same rules. They have the same conjugations and the same patterns – with the exception of changes triggered by one or several glides in the root:

Bare form:	$C_1aC_2C_3aC_4$,	$iC_1aC_2C_3aC_4$	participle	$mC_1aC_2C_3aC_4$
Reflexive form:	$tC_1aC_2C_3aC_4$,	$yətC_1aC_2C_3aC_4$	participle	$mətC_1aC_2C_3aC_4$
Passive form:	$uC_1aC_2C_3aC_4$,	$yuC_1aC_2C_3aC_4$	participle	$muC_1aC_2C_3aC_4$

On quadriliteral verbs, there are a few pioneering studies such as Kamil (1963), and there has been some more recent interest, both for Classical Arabic (Procházka 1993) and dialects (Madouni 2001; Holes 2004; Albader 2016), but on the whole the area has been little studied.[9]

Personally, since I began taking an interest in quadriliterals in Ḥassāniyya, I have collected a corpus of approximately 800 distinct roots, which are represented either by a verb (in their bare and/or reflexive form) or by a participle.[10] In addition to my personal data, both published and unpublished, which constitute the major part of my corpus, I further combed through the dictionary compiled by Heath (2004 – henceforth ʜᴇᴀᴛʜ) on Mali Ḥassāniyya and the lexicon compiled by Tauzin (1993).

Based on this corpus, I have already carried out a study (Taine-Cheikh, forthcoming) on the semantics of reduplicated biliteral roots (118 roots of the type $C_1C_2C_1C_2$ or, to simplify, 1212).

In the framework of the present study, I have retained 148 quadriliteral roots susceptible of being related to Berber. The verbs are divided into three groups, depending on the depth and likelihood of their relation to Berber. We will first examine those with nearly certain Berber origins (Section 2). This will be followed by those where Berber origins are probable (Section 3), and we will end with those where the etymology is more complex (Section 4).

Remarks on presentation:

–   Entries follow the root. Data preceding the triangle are from Ḥassāniyya. Following data refer to other dialects (whether Arabic or Berber).
–   When several different meanings are found in Ḥassāniyya, a number is used to distinguish them, in reference to the *Dictionnaire* (Taine-Cheikh 1988–1998).

---

9.   In the Western Algerian dialect she studied, Madouni noted 148 quadriliteral verbs (for 102 quadriliteral roots) as compared to 1767 verbs with triliteral roots (Madouni 2003: pp. 243–4).

10.   The presence of lone nominal forms makes determining the root more haphazard, especially in the case of borrowings from Berber where vowel length often varies and the nature of certain consonants (affix or root) is not always clear.

- An asterisk ⋆ preceding a Ḥassāniyya form indicates that I found the form in the literature but have been unable to validate it.
- The following abbreviations are used:

adj.	adjective	PL	plural
Berb	Berber	PM	pan-Maghreb
Cl.	Classical Arabic	prep.	preposition
F	feminine	region.	regional
G	Tuareg Kel-Gərəs	SG	singular
Ghad	Ghadamsi	Tach	Tashelḥît
Hass	Ḥassāniyya/hassanized	Tam	Tamazight
K	Kabyl	Tu	Tuareg
M	masculine	unus.	unusual
MA	Moroccan Arabic	v.	verb
masd.	maṣdar (action noun)	var.	variant
n.	noun	W	Tuareg tawelləmət
part.	participle	Y	Tuareg tayəṛt
		Zen	Zenaga

- Works that are cited often are referred to by their author's name in small capital letters, with the exception of Zenaga (Zen) which, in the absence of any further specification, refers to my *Dictionnaire Zénaga–Français* (Taine-Cheikh 2008c).

2. Probable borrowings from Berber

In the following examples, convergence is not always complete, but is sufficient on the whole to make borrowing the most plausible hypothesis.

2.1 Strong formal and semantic convergence

2.1.1 *With no other known origins*

1. ẒWẒY *ẓẓawẓa* 'show disappointment with what one has received'. ▲ Tu PRASSE ^{WY}*əẓwəy/zəẓwəy* 'welcome reservedly'; ^W*uẓaẓ* 'deem insufficient (in quantity or quality)'.
2. ZRWL *mẓaṛwäl*, part.-adj. (v. rare) 'odd-eyed' – var. of *ẓəṛwâl*. ▲ Tam TAIFI *aẓerwal* 'who is cross-eyed, who squints = who has green eyes'.
3. ZLMṬ *ẓalmaṭ* 2. 'deviate (as in one's gaze)'. ▲ Berb ZLMḌ: TAIFI *zelmeḍ* 'be left-handed; be deformed; be clumsy'; Zen (with ⋆L > Y) *žäymuḏ̣* 'left-handed'.
4. ZWLY *ẓawlä* 1. 'form, lead a caravan of camels (*aẓālây*)'. ▲ Tu PRASSE ^{WY}*əẓləy* 'set aside; match (merchandise)'; *aẓälay* 'salt caravan'.

5. VRKK *värkäk* (region.) 1. 'make fall noisily, forcefully'. ▲ Berb: ᴛᴀɪꜰɪ *ferkek* 'crack, crumble'; ᴅᴀʟʟᴇᴛ *fferkekk* same meaning; ꜰᴏᴜᴄᴀᴜʟᴅ *ferekket* 'be open'.

6. GRWY *gärwä* (East, var. *šowṃä*) 'put on a camel crupper (*ägärwä*)'; *ägərwi* 'body of the *täzâyä*, rectangular bag made out of decorated leather'. ▲ Tu ᴘʀᴀꜱꜱᴇ ᴳ*egärwäy* 'crupper rope (which attaches the camel's saddle to its tail)'; ᵞ*egärwäy* 'ornamented skin pouch'.

7. GYMR *gäymär* 'far away hunt'; *gäymârä* 'far away hunting'. ▲ Berb 'fish; hunt': ɴᴀïᴛ-ᴢᴇʀʀᴀᴅ GMR(3).

8. ɣWBY *ɣawbä* 1. 'halter (a camel)'; *aɣâbä* 'chin strap (to halter a camel)'. ▲ Zen *yäšquḅä* 'tie a rope to the lower jawbone (*aɣabä*)'; Tu *aɣaba* 'bit'.

9. GNDZ *mugändäz* '(animal) which has had its four legs broken to prevent it from escaping'; *gändûz* 1. 'middle part of the leg (sheep)'. ▲ Zen *ägundiž* 'shin' (ʙᴀꜱꜱᴇᴛ *agouandouž* 'calf'). Berb: ɴᴀïᴛ-ᴢᴇʀʀᴀᴅ GNDZ 'calf'.[11]

10. ZKRN *ẓakṛan* 'lock up'; *ẓakṛûn* 'modern, imported lock'. ▲ MA ʙᴇᴀᴜꜱꜱɪᴇʀ *z.kr.m* 'lock'; *zakrū(m)* 'lock'. Berb ᴛᴀɪꜰɪ *azekrum*, ᴅᴀʟʟᴇᴛ *azekṛun* 'lock'.

11. KMBR *kämbär* 1. 'make, bear or use the mark *äkämbûr* (fold of skin which forms on the nostrils of cows after vaccination)'. ▲ Tu ꜰᴏᴜᴄᴀᴜʟᴅ *ăkenbour* 'small outgrowth of skin on the nose …'; ᴘʀᴀꜱꜱᴇ *akănbor* 'ornamental mark …'.

12. ɣRĞY *ɣäršä* (region.) 'zigzag, walk like a fennec'; *aɣərši* 'fennec'. ▲ Zen *äɣəršäy* 'fennec'; ꜰᴏᴜᴄᴀᴜʟᴅ *äẖorhi* same meaning.

13. VRKS *värkäs* 1. 'hatch (eggs); break eggs (hen)'. 2. 'squirm, fight' + VRKṢ *faṛkaṣ* (Mali, ʜᴇᴀᴛʜ) 'squirm, fight'. ▲ Berb: cf. ᴅᴀʟʟᴇᴛ *fferkes* 'be deteriorated'. ᴛᴀɪꜰɪ *ferkeš/feršeš* 'crack, crumble'. MA ᴄᴏʟɪɴ *fərkəs* 'frolic, squirm'.[12]

14. ŽYKR *žäykär* 1. 'beat down (gum)'. 2. 'braid in two strands'. *(ä)žäykâr* 1. 'pole for gathering gum'. ▲ Zen *ižī'gär* 'rope for attaching loads', *äža'gäri* 'pole for gathering gum'. Berb ZKR, cf. K ᴅᴀʟʟᴇᴛ *iziker* 'rope'.

15. RYVY *mräyvi* part.-adj. (v. rare) 'from the harmattan'; *irîvi* 'harmattan'. ▲ Zen ɴɪᴄᴏʟᴀꜱ (p. 442) *irīfi* same meaning. Tam ᴛᴀɪꜰɪ *raf* (rare) 'be very thirsty', *irifi* 'thirst, great thirst'.[13]

---

11. Cf. Monteil (1952: p. 114).

12. Behnstedt & Woidich liken FRKS/Ṣ to FQS/Ṣ: cf. *faqasa* 'break, destroy (an egg, for a bird)' in Classical Arabic.

13. Cf. AM ʟᴏᴜʙɪɢɴᴀᴄ *rīfî* 'hot southerly wind (< Berb *irifi*)', *rwfy* 'suffer from the heat, suffer a heatwave'.

16. VSKY *väskä* 'spend the *tivəski*', *tivəski* 'beginning of the dry season (intermediate season: spring and/or autumn)'. ▲Zen *yäväskä* 'spend the spring'. 'Spring': Zen *tfəskih*, Tu *tafsît*, K *tafsut* …[14]

17. GMKY *gämkä* (S.-W. var. of *dämkä*) 'learn to read, mumble'. ▲Zen *yägämkäh/yäžgämkäh* 'spell out' < GMG *yugṃäg* 'follow' (Berb GMK).

18. RWGŽ *rowgäz* (S.-W.) 'go on foot'; *ärägâz* 'human being, individual, person'. ▲Zen *ärägäz̲* 'person'. Tu FOUCAULD *regeh* 'march in step'; *aragah* 'group of people on foot marching in step'.

19. LYWŠ *läywäš* 'cover with a *ilīwîš* (sheepskin with its fleece)'. ▲Zen (with *L > Y) *iyīš* 'prayer rug (a single skin)', *iyäwšän* 'large blanket'. K DALLET *els* 'be dressed'; *ilis* 'sheared fleece'.

20. Ɣ/QMBR *ɣambär* (*qambär* in the East) 'envelop someone (in or with)'. *tɣambär*, masd. *ɣambûr/taqambūrît,* 'bundle up'. ▲Ghad LANFRY *gumber* 'bundle up, for protection from the cold'.

21. GYLL *gäyläl* 'cut, shorten the tail (animal)'; *ägīlâl* 'which has a shortened tail'. ▲PRASSE [WY]*giləl* 'have a short tail'. Zen *gäyyiy* 'which has a short or shortened tail'.

22. WYTL *wäytäl* 'attach the *īwâtlən* (part of the well rope which goes around the belly of the animal pulling the rope)'. ▲Zen (with *L > Y) *uwädiyän* 'well rope which is attached around the animal'. Berb NAÏT-ZERRAD DL(2) 'braid'.

23. DVRY *dävrä* 1. 'put on the saddle rug'. ▲Tu PRASSE [W]*səddəfurət* 'put an *adäfor* as a saddle on (a mount)'. Zen *tädäffurt*.

24. GNZY *tganza* (Mali, HEATH) 'circulate'; *tāgənzä* 'wooden circle of the delou (skin bag used to draw well water)'. ▲Zen *tägänžäh* same meaning. Tu PRASSE [W]*gänzäw* 'be arched', [WY]*täganze* 'bow, arch; *âga* hoop (made of iron)'.

25. DRṢY *daṛsa* 'have dactylitis'. *âdəṛṣ/ädəṛṣ* (or *ḍ*) 'dactylitis'. ▲Tu FOUCAULD *aḍräz̲* 'dactylitis'.

26. TWQY *towqa* 'be attacked by the *tâqa* insect (palm tree)'/TWKY **mtowki* 'affected by the *tâkka* disease (dates)'; **tākkä* 'disease which prevents dates from ripening (because of the wind which covers them in dust)'. ▲To FOUCAULD *tahoḳḳa* 'dust'.

27. ŽYWY *mžîwi* 'suffering from *ižîwi* (illness among women due to air drafts)'. ▲Zen *iʾž̲əwi* 'wind'. Ouargli DELHEURE ẒW *taẓwət* 'small wind, petit vent, breeze'.

---

14. See Taine-Cheikh (2015) on the season 'spring' – the etymological relations with the verbs 'disintegrate, come undone' in northern Berber and the seemingly secondary relation to the term for feast (Berb *t(a)faska* < Latin *pascha* < Hebrew).

28. ZNGY *z̧angä* 1. 'transform into a tributary'. 2. 'act meanly towards someone'. ▲ Zen *az̧nug* pl. *uz̧nägän* 'Zenaga, tributary' (meaning specific to Mauritania).[15]

29. WLTM *wältäm* 'make undergo (she-camel) a forced lactation operation (by inserting the fist in her uterus)'; *towlǝtmît* 'fist; punch; (+ prep. *mǝn*) fistful of'. ▲ Zen *täwlluṃt* 'punch' (BASSET *touletemit*). Tach DESTAING *īmi ntu̧eltīmt* 'wrist'. Beni-Snous DESTAING *θižeḷθemθ* and K DALLET *igeltem tigeltemt* 'biceps'.

30. ƔZMR *ɣazmär* 'be hidden by one's beard, hair'. ▲ Zen *yaɣazmäräʰ* same meaning and (with *Ɣ > ʔ) *taʔMärt* 'chin'. K DALLET *aɣʷesmaṛ* 'jawbone'; Ghad LANFRY *ɣusmar* 'chin', *aɣʷesmaṛ* 'jawbone'.

31. RWGL *rowgäl* 1. 'trot the pace'; *rowgâlä* 1. 'pace (horse, camel)'. ▲ Tu FOU-CAULD *ergel* 'close' and *regiregi* 'pace gait'.

32. VLWY *mǝtfalwi* (Mali, HEATH) '(well) arranged (explanation, etc.)'. *tāvǝlwît/tīvǝlwît* 1. (rare) 'door with two panels'. ▲ Tu FOUCAULD *taflout* 'wing (of a window)'; PRASSE [Y]*tǝfǝlwit* 'door'; [Y]*efliwǝs* 'disorder'.[16]

2.1.2   *Foreign terms found in Berber*

In at least two cases, the lexemes found in Berber were first borrowed from other languages.

33. GWDR *gowdär* 'surround'; *ägādîr* 'stone enclosure'. ▲ Tu FOUCAULD *agadir* 'wall'. NAIT-ZERRAD GDR(1): of Punic origin (cf. Laoust 1920: p. 3; Vycichl 2005: p. 3).

34. BRGN *bärgän* 1. 'lower the tent'. ▲ Colin (1926: p. 58): Latin *barca* > Andalusian *barga* and Berb *tabergent* 'grinder; hut; hair tent'. FOUCAULD *ăberġen* 'hair tent (which is not found in Hoggar)'.

2.2   Partial formal and/or semantic convergence

2.2.1   *Partial semantic convergence*

In the following four examples, the Ḥassaniyya verb is formally similar to the Berber but its meaning is probably derived, rendered more specific.

35. ZWGR *mzowgär* 'immunized (against an illness)' – var. of *zâgǝr*. ▲ Zen *yǝzgär* 'come out; come out of the ground, sprout'.

---

15. Those who speak Zenaga tend to not see themselves as Berber, nor even Ṣanhāǧa, and even less as *Znâgä* since this latter term now denotes a despised segment of society. On the change *Iz̧nāgǝn > Ṣanhāǧa*, cf. Colin (1930).

16. 'Disorder'/'(well) arranged (explanation, etc.)': perhaps the outcome of antonymous semantic evolution.

36. NKRY *näkrä* 'return (for an illness); resume (for a plant)'. ▲Zen *yunkär* 'rise'. Berb NKR same meaning.

37. NGRY *nägrä*, masd. *änəgri/āmnägri*, 'tune an instrument depending on the mode'. ▲Zen *yängārä* 'start in reverse, go in the opposite direction'; * əngīri* 'crossing …'. Tach DESTAING *ngiri* 'separate'.[17]

38. RNKY *ränkä* 'mistreat a slave'; *ärənk* 'bad, lazy slave'. ▲DALLET RNK *ṛennek* 'disturb; bother; annoy'.

2.2.2   *Partially divergent roots*

These quadriliterals in Ḥassāniyya show at least one more root consonant than the Berber lexemes they are comparable to. In the first three examples the consonant is a nasal, a liquid or both. In the following three, it is a word initial stop – a more unusual phenomenon, especially in the case of K'.[18]

39. ZWRN *mzowrän* 1. 'which does not drink properly, does not fill up properly (camel)'; 'palm rib used for making baskets'. ▲Zen *yəzrāh* 'have an empty stomach'. Tu PRASSE ^WY*azru* 'be listless, very tired (ill)'.

40. GŽML *gäžmäl* 'shorten'. *mgäžmäl* 1. 'short (hands or feet)'. ▲Berb (with contamination across 2 roots?): cf. Zen *yugžäm* 'shorten' and Berb GZL, cf. NAÏT-ZERRAD.

41. RGŽN *rägžän* 1. 'make a recalcitrant she-camel kneel (rutting camel)'. ▲Zen *yugän* 'be kneeling, make kneel (camel)'; *yäžgän* 'make kneel (camel)'. Berb 'sleep; kneel': NAÏT-ZERRAD GN(5).

42. KRẒY *kaṛza* 'mount (young untrained animal)'; *ākəṛzi* 'young camel being trained'. ▲Berb RẒ/ṚẒ 'break', cf. Zen *yaṛza* 'be broken; break'; ouargli DEL-HEURE *eṛẓ*; etc.

+

43. ṬRẒY *ṭaṛza* (S.-W.) 'put a *ṭäyṛza* (a rope for training) on a bovid'. ▲Zen *tīṛẓah* same meaning < *yaṛẓa* (RẒʔ) 'break'.

44. TYRS **täyräs* 'dig a new well'; **tīr(i)s* 'freshly dug well'. ▲Tu FOUCAULD *îres* 'well (hole dug in the ground to draw water, > 2 m deep)'.

---

17.   These forms could be related to roots without N such as K DALLET *egri* and Ghad LANFRY *eğr* 'warp a thread (weaving)'. In which case the N would have affixal origins.

18.   The initial coronal consonant could come from the noun prefix *t-* (contextually emphatisized in the case of Ṭ). I consider that the K is a variant of Ṭ but one could also consider it a variant of Q (on this preformant and its variants, cf. Lentin 2010).

2.3 Roots where the 1st consonant was originally an affix

2.3.1 *With the factitive affix* 's'

Quadriliteral verbs in Ḥassāniyya often have more specific meaning than the Berber verbs. Moreover, in this latter case, one must posit a change in the liquid (R < L).

45. S/ṢYVṬ *säyvaṭ/ṣäyvaṭ* 1. 'say goodbye'; masd. *ṭsäyvîṭ/tāmṣivâṭ/tīmṣivâṭ*. ▲MA *sifəṭ* < Berb FḌ (Pellat 1950; Kossmann 2013: p. 187). NAÏT-ZERRAD FḌ(1) 'send; say farewell …'. Zen ʔFḌ *ṣuʔf(f)uḍ* 'accompany'.

46. ZWZL *zowzäl* 'castrate (camel, bovid, human)'; *äzūzâl* 'castrated (bull) camel, gelding'. ▲Zen (with *L > Y) *äẕuẕäy* same meaning. DALLET *iwzil* 'be short', *ssiwzel/zziwzel* 'shorten'. COLIN *zūzəl* 'castrate by ablation', *cp.* Berb *zowzəl* 'shorten'. Monteil (1952: p. 118 n. 72): PM < Berb.

47. ŠWNN *šownän* 'begin the training (of a camel)'. *äšänân* 'young saddle camel in training'. ▲Tu FOUCAULD *anen* 'be trained', *sinen* 'train'. Zen *aʔšänän* 'camel or horse having started training'.

48. ŠWMY *šowmä* 'put on a camel a *äšâmä*'. *äšâmä* 'strap going from the hind part of the belly to in front of the sheath (to hold the saddle in place)'. ▲Berb ɣM. Zen (with *ɣ > ʔ) *yäššäʔmä* 'make sit', factitive form of *yaʔmä* 'sit'.

49. ŠYḌḌ *šäyḍaḍ* 1. 'suckle a female other than one's mother (kid, lamb)'. *äšäyḍâḍ* 1. 'who suckles …'. ▲Berb: Zen ḌḌ *yuḍḍaḍ* 'be suckled; suckle', *yaṣṣuḍaḍ* 'nurse'. FOUCAULD *eḍḍeḍ* …

50. ŠKRV *šäkräv* 'immobilize (an animal) in kneeling position using tethers'; *ššäkräv* 'curl up'. ▲Berb FOUCAULD *ekref* 'hobble (the two front legs)'; DALLET *ekref* 'be paralyzed'; Zen *yugräf* 'retract'.

51. ŠKRṬ *šäkraṭ* 1. 'torture, morally wound'. ▲Berb KRḌ: PRASSE sᵂᵞ*əkraḍ* 'scrape'. Zen GRḌ *yugraḍ* 'scrape (to clean …)'.

52. SGRY *sägrä* 1. masd. *tāmsəgrît*, 'refuse someone something'. 2. masd. *āmsəgri*, 'apologize'. ▲Zen *yugräh* 'hear'; *yässəgräh* 'apologize (for something and/or to someone)'. Berb NAÏT-ZERRAD GRH 'discern, be careful of'.

53. SDBY *sädbä* 'make leave in the afternoon'; *ssädbä* 'leave in the afternoon'. *tāsədbît* 'afternoon departure'. ▲FOUCAULD *adou* 'arrive in the afternoon at', *sidou* 'make go in the afternoon in'. Zen *təđäbbäđ* 'departure in the afternoon', *yässədbäh* 'make leave the afternoon'.

54. SNTY *säntä* 'begin'. ▲Zen *yəntä* 'sting'; *yässäntä* 'begin'. FOUCAULD *ent* 'be begun; be solidly fixed'; *sent* 'begin'.

55. ŽWŽY *žowžä* (rare) 'dig'; *mžowži* 'which produces a strangled sound, as if hoarse'. ▲Berb ɣZ. Zen (with *ɣ > ʔ) ʔŽ *yaʔẕ* 'dig (hole)', *äẕoʔẕ/äẕäwẕ* 'place where water may be found'.

56. ŠKML *ššäkmäl* 'tighten; shrink; shrivel up'. ▲K DALLET *ekʷmer* 'be narrow'.

2.3.2   *With the nasal affix* m-/n-

Contrary to the prefix *m-* which is used in both Arabic and Berber, *n-* is a noun prefix only in Berber.

57. MGRD *mmägräd* 'suffering from *tmägrîd* (camel)'. *tmägrîd* 'ruptured tendons of the cervical ligament'. ▲ Zen *əgaṛḍ* 'neck'. PRASSE ^{WY}*egäräḍ*/^W*temäggäräṭ* 'nape'.

58. MTRG *mäträg* 'attach the strap *tīmətrâg* made of braided leather which goes around the pack saddle (on a women's camel saddle)''. ▲ Zen *tämätrəkt* same meaning. FOUCAULD TRG *etreġ* 'free completely', *émetreġ* 'chain going around women's mehari saddles', *témetrek* 'cord made of braided leather'.

59. NYWL *näywäl* 'lead an animal drawing water from a well'. *änäywâl* 1. 'leader of the animal (trained) for drawing water'. ▲ Tu PRASSE ^{WY}*əwəl* 'turn'. Zen *änäwäy* same meaning than *änäywâl*.

3.   Possible borrowings from Berber

3.1   Cognate(s) in a single Berber dialect

In the following examples, borrowing from Berber is plausible. Moreover, several elements point at times to double interference, where the target language (Ḥassāniyya) 'gives' a verb form to the primary source language or influences it.

     All correspondences concern Zenaga, except for *maqras* (in 3.1.3) which was observed in Mali by Heath.

3.1.1   *Nominal cognates*

The verbs (or participles) are generally formed on a nominal base (noun or adjective) having an equivalent in Zenaga with highly similar (if not identical) form and meaning. At times however, the Ḥassāniyya verb form in appears to be directly formed on the Zenaga nominal form. Verbs and nouns either share the same root (e.g. ŠGDL) or not (e.g. GWṬY/GWŽ, MZRY/ZRY).

60. GLMN *gälmän* 1. 'winnow (millet)'. 2. 'reduce (skin)'. *ägälmûn* 'germ (cereals)'. ▲ Zen *əgäyuṃän* 'millet husk, germ'.

61. GWṬY *gowṭa/tgowṭa* 'have a sore spine'. *ägowž* 'spine, spinal column'. ▲ Zen *ägäwẓ* 'wide dune'.[19]

---

19. *aɣawž* 'spinal column' and *aɣawẓ* 'dune' share the idea of a ridgeline.

62. GWBN *tgowḅän* (Mali, HEATH) 'behave like a glutton'. *gābûn* (without article) 'spotted hyena'. ▲Zen *gähūh* 'spotted hyena' (< Wolof *gakh* 'howl').

63. ŠGDL *šägdäl* 'serve as the basis for an anvil; write on the back of …'. *äšəgdäl* 'anvil base, support'. ▲Zen *äšugdäy* 'desk blotter'.

64. ŠWGY *šowgä* (S.-W./*taʕwä* in East) 'catch/give whooping cough'. *täšâgä* 'whooping cough'. ▲Zen *taʔšägäh* same meaning.

65. ƔWŽY *ɣawžä* 1. 'dig a *(a)ɣawži* (a ditch around a tent)'. ▲ Zen *äɣaẓäh* 'ditch (to channel rain water away from houses)'.[20]

66. MZRY *tmäzrä* 1. 'stand aside (a little)'. *âzri* 'side'. ▲Zen *əzri* same meaning.

67. BWLY *tbowlä* 1. 'become an *äbūlây*'. *äbūlây* 'a plump, healthy, very young animal'. ▲Zen (with *L > Y) *äbuyäh* same meaning.

68. GṢMṬ *mgaṣmaṭ* 'having very small ears'. ▲ Zen *gaṣmuḍ* 'narrow, tight (bed, clothes); having small ears'.

69. MYGN *mmäygän* 'which still has a taste of tannin (waterskin)'. *imigîn* 'taste of tannin'. ▲Zen *əmiʔgən* same meaning.

70. NYTY *əmnäyti* 'covered with *inîti*'. *inîti* 'cram-cram, *Cenchrus biflorus*'. ▲Zen *ənəḏīh* same meaning.

71. VŠKY *tväškä* 'disperse (clouds …)'. ▲Zen NICOLAS *təfəššəkť* 'empty area (without trees nor tents)'.[21]

3.1.2   *Pluralities of cognates*

Zenaga appears to be the source from which Ḥassāniyya borrows what is often a nominal form. In the three latter cases, the verb form in Zenaga, of the type *ya-Ca(Ca …)Ca(h)*, further appears to be influenced by the Arabic form.

72. GṚGṬ *gaṛgaṭ* 1. 'add too much salt'. 2. *'drink the potion called *girigṭa* (which is very salty)'. ▲Zen *gaṛguḍ* 'too salty', *yägguṛgaḍ* 'become too salty'.

73. MRKY *tmärkä* 'cease giving milk (dairy animal); roam off (cattle)'. *tīmərkît* 'milk cow which is not in the lactation period and has not been impregnated'. ▲Zen *tämmərkäh* 'no longer have milk and be free to roam'. *tmərkiʔḏ/ tiʔmərkiʔḏ* 'milk cow which is not in the lactation period'.[22]

---

20. *aɣaẓäh* is maybe an irregular diminutive form derived from the same root as *yaʔž* 'dig'.

21. On possible relations between the notions of dispersal (*tväškä*) and springtime (*tivəski,* cf. 16), see Taine-Cheikh (2015).

22. If, as I believe, *îrki* PL *ârkän* 'calf' belongs to the same word family, then there is certainly a cognate in Tuareg: FOUCAULD *éberkaou* 'non-weaned calf, suckling calf which has begun eating grass'.

74. GRWL *gärwäl* 1. 'encase a well in its submerged part'. *ägerwâl* 'large quantity of liquid'. ▲Zen (with *L > Y) *yäžgärwäy* 'encase a well in its submerged part'. *ägərwäy* 'encasing in the groundwater'.

75. TYMŠ *täymäš* 'be, become anemic'. *tîmši* 'illness due to the absence (of a sort) of milk'. ▲Zen *tīmših* same meaning; *yättäymäššä* (of Hass form) 'be ill from lack of a certain milk'.

76. GNDY *gändä* 'catch the illness attributed to excess consumption of something (tea, tobacco, etc.)'. ▲Zen *igəndih* 'food poisoning'; *yägändäh* (of Hass form) 'suffer from food poisoning'.

77. BWRY *bowrä* (*b/ḅ*) 'suffer from *tābūrît*'. *tābūrît* 'sleeping sickness'. ▲Zen *täburiˀđ* same meaning; *yäbäwräh* (of Hass form) same meaning.

3.1.3   *Verbal cognates*

In the following examples, it is a verb form which goes from one language to another (possibly twice – which would explain why there are two Zenaga forms for the first two verbs). In *maɣnä* and *šänkär*, the prefix consonants in the source language have become root consonants.

78. ŠYVR *šäyvär*, masd. *täšävârət*, 'be from the same camp as someone', 'be the neighbor of'. ▲Zen *yäššäffär* 'be the neighbor of' – var. (of Hass form) *yäššäyväräh*.

79. GNGL *gängäl* 1. 'become hairy (plant)'. 2. 'cause itchiness, itch'. ▲Zen *yäggungäy* 'itch (for the skin)' – var. (of Hass form) *yägəngäyä*.

80. MQRS *maq̇ras* (Mali, ʜᴇᴀᴛʜ) 'be ill from lack of an usual food (milk, rice, etc.)'. ▲To Mali (Heath 2006) *-əmməɣræs-* same meaning.

81. MɄNY *maɣnä* 'become angry'. ▲Zen *yäṃṃuɣnāh* 'become angry' – a (reflexive) form derived from *yuɣnāh* 'make angry'.

82. ŠNKR *šänkär* 1. 'scrape (a bone) – var. de *šäkkär*. ▲Zen *yäššənkär* factitive form of *yunkär* 'be scraped (bone)'.

83. ŽWNY *žownä* 'butt someone with the horns'. ▲Zen *yäžīnä* 'give a butt (for a bovid)'.

84. NYMŠ *näymäš* 'show off'; *tnäymäš* 'look at with curiosity'; 'admirable'. ▲Zen *yänmäššä/yärmäššä* 'behave kindly towards'.

85. NKFR *näkfär* 'promise (v.)'. *tānəkvârət* 'promise (n.)'. ▲Zen *yäšnäkfär* 'promise (v.)'.

3.2   Presence in Ḥassāniyya of a distinctive 'Berber' feature

In the following cases, nothing has been found in Berber to corroborate the hypothesis of borrowing. It is thus highly conjectural.

3.2.1   *Nouns with unusual syllable structures and affixes*

86. MNDY *mändä* 'suffer from *iməndi*'. *iməndi* 'camel illness due to eating grass which is too wet'.

87. ŽɣND *žaɣnäd* 'contract the respiratory tract illness (animals, pej./humans)'. *žäɣändi* (S.-W., without article) 'respiratory tract illness (animals, pej./ humans)'.

88. GWGM *mgowgäm* 'suffer from *tāgūgâməṭ*'. *tāgūgâməṭ* 'trypanosomiasis (of the camel)'.

89. BWRŽ *mətbowräž* 'lacking resistance because untrained'. *äbowräž* 'young camel lacking in resistance'.

90. GWŠṬ *gowšaṭ* 'make rapid and unexpected turns'. *ägušâṭ* 'which has small pointy ears (horse …); who has small ears (human)'.

91. KWMR *kowmär* 'castrate (donkeys, horses)'. *äkūmâr* 'gelding (horse)'.

92. MWLY *mətmowli* part.-adj./*mowli* PL *imūlyân* 'young and promising (camel)'.

93. MWKY *mowkä* 'roam freely, not be kept at camp (for bovids, especially bulls)'. *äṃäkây* PL *āṃṃakây* 'group of animals generally having no calves and roaming freely'.

94. NZGR *näzgär* (rare) 'develop an *änäzgâr*'. *änäzgâr* 'sore on the backbone (camel)'.

95. GNGY *gängä* 'sift, winnow'. *(ä)gängây* PL *āgnâgä* 'sieve, screen'.

96. ŠWTY *šowtä* 'throw far away'. **äšowtāy* PL *ašwātāy* 'lever'.

97. SRBT *särbät* 1. 'sift with a coarse mesh sieve'. 2. 'gobble up quickly'. *(ä)särbât* PL *āsrābît* 1. 'large mesh sieve'. 2. 'rapid swallowing'.

98. NWDR *nowdär* 'train (horse)'. *änowdâr* 'cross bar (of a pendulum well)'.

99. BWKK *mbowkäk* 'resembling *äḅäkâk*'. *äḅäkâk* 'seyal acacia resin (false gum)'.

100. RWDN *rowdän* 'rain for a long time, in a fine shower'. *(ä)rädânä* 'winter drizzle'.

101. BMBY *bämbä (t-tṛâb)* 'make a pile of (dirt, sand)'. *ibämbi/äbämbä* 1. (rare) 'small mound (of sand, dirt)'. 2. (East, Néma) 'bench made of banco, stone'.

102. TKRY *täkrä* (S.-W./*täkrär* in East) 'put a piece of cloth on one's head to carry a load'. masd. *ätäkrä*.

103. NKMṬ/Ḍ *tnäkmaṭ/tnäkmaḍ* 'contract (from the heat)'. *änəkmâṭ* 'shriveled up date, juiceless'.[23]

104. ŽWRV *žowräv* 'eat *täžârəvṭ*'. *täžârəvt* 'millet porridge with a lot of milk'.

---

23. The alternation *ṭ/ḍ* is an additional indication of borrowing, on one hand because *ḍ* is much more frequent in Zenaga than *ṭ*, and on the other hand because variation tends to be more frequent with borrowings.

105. VNGR *vängär* (*v/ɣ*) 1. 'husk'. 2. 'make an incision in the skin to remove a thorn' – var. of *mängär*. *tīvəngrît* 'inside of the baobab fruit (powdered)'.
106. MDRY *mädrä* 'pull the upper threads with a *mədrä* (to tighten the weaving)'. *täṃädrît* (rare, Adrar) 'instrument for cutting palm shoots'.
107. ŽNKṬ *žänkaṭ* 1. 'trip someone (wrestling)'. 2. (among the Nmâdi hunters) 'cut the hamstrings of prey to prevent them from escaping'. **tāžänkaṭ* (among the Nmâdi] 'foot muscle, hamstring of prey'.
108. ŠRTT *šärtät* 'leave the *Idowˁîš äḅäkâk* in dissent for their rivals, the faction of the *Šrātîî*. *šärtât* (without article, East) PL *šrātît* 'hyena'.[24]
109. BWST *bowsät* 'fire brand with the mark of the Ideybūsâî'. *Idäybūsât*: name of a Mauritanian tribe.[25]

3.2.2 *Presence of non-Arabic phonemes:* ẓ, dʸ,tʸ

110. ND^{Y}WR *nädʸwär* (var. *kädʸwär*) 'search for a rare product or seek to replace it with a similar product'. *tnädʸwär* 'cave in (well)'.
111. KWT^{Y}M *kowtʸäm* 'punch, box'.[26]

4. Puzzling etymologies

4.1 Insufficient or contradictory indications

4.1.1 *Zenaga verb of the type* yaCa(Ca …)Cah

Such a verb, in the absence of other indications, would seem to point to Ḥassāniyya as the source language. The following six verbs are therefore probably of Arabic origin. One should note the retention of ɣ (unusual in Zenaga) in three of the verbs, as well as the fact that the last two verbs are also found in Maghreb Arabic.

112. WNGL *wängäl* 'take turns slitting the throat (of animals)'. ▲Zen *yäwängäy* same meaning.
113. GRBZ *mgärbäz* 'which has a middling belly'. *tgärbîz* 'state of a camel with a middling belly'. ▲Zen *yägärbäzzäh* 'take on a middling belly (camel)'.

---

24. The meaning 'hyena' could be secondary in regard to the use of the plural as a proper noun. It should be noted that the Idoˁîš are one of the rare tribes which, in recent times, still recognized their Berber origins (Ṣanhāja) – evidenced moreover by the form of their name.

25. Tribal name which also sounds Berber.

26. See Hass. *kätʸtʸ* 'give a blow, hit (punch, stab with a sharp object …)'.

114. ɣRNG *mɣarnäg* 'which has deep set eyes'. *ɣaṛnûg* 'cavity, indent, socket'. ▲Zen *yaɣarnägäʰ* 'be deep set (eye)'.

115. QRWṬ *qaṛwaṭ/tqaṛwaṭ* 'rumble (guts, belly)'. ▲Zen *yaqarwäṭaʰ* same meaning.

116. ɣZ ɣZ *ɣazɣaz* 'make grit (teeth, leather)'. ▲Zen *yaɣazɣazza* same meaning. PM BEAUSSIER *ɣazɣaz* 'grit under the teeth; grind the teeth'.

117. ɣYDN *ɣaydän* 'wean too early (lamb, kid)'. ▲Zen *yaɣayðänäʰ* 'be weaned'. MA LOUBIGNAC *ɣydn* 'separate lambs or kids from their mothers, preferably during the day, to prevent them from suckling'.

4.1.2   *Possible internal changes*

Despite the presence of 'Berber' noun prefixes, the semantics of the triliteral and quadriliteral roots makes the hypothesis of changes internal to Arabic plausible, either through addition of a glide, or transformation of the affix *m* into a root consonant. In the first case however this entails positing metathesis (as well as emphasis of the Z).

118. ẒYWN *məẓẓäywän* 'discerning music lover'. *aẓāwân* 'concert of Moorish music'. ▲Guignard (2005: p. 28 n. 1): < Arabic root WZN 'weigh; measure'.

119. VWŽR *vowžär* 'still hungry after just having left the trough'. masd. *ävowžâr*. ▲[Cl. FŽR] Hass *väžžaṛ* 1. 'leave at dawn …'.

120. MRSL *märsäl* 'bring (animal) to a salt lick'. *ämərsâl* 'salty earth for cattle'. ▲[Cl. RSL] Hass *râsəl* 'release cattle in successive waves (in particular towards a trough and salt)'.

121. MŽLY *tmäžlä* 'reach the age of *muẓli*; be fit (like a *muẓli* cow)'. *muẓli* 'heifer of approximately four years which has not yet calved, for which gestation has "dissipated"'. *âmžəl* PL *īmžəllâ* 'bull'. ▲[Cl. ŽLY] Hass *žlä* 1. + prep. ʿ*an* 'remove from; dissipate (sadness, worry, fatigue, etc.)'. 2. '(make) lose'.

4.2   Notable correspondences outside of Berber

Various clues seem to indicate borrowing from Berber. However, correspondences with other languages instill reasonable doubt.

4.2.1   *In Maghreb Arabic*

There is very little doubt as to the Berber origins for the first two, somewhat more doubt for the following ones, especially the last two. There are still strong formal and semantic correspondences, but Arabic could be the source language.

122. BRKŠ *bärkäš* 'moisten couscous etc. with (milk, water)'. **bərkūkīš* 'couscous ball'. ▲Berb ɴᴀɪᴛ-ᴢᴇʀʀᴀᴅ BRKS(2): ᴅᴀʟʟᴇᴛ *bberkukes* 'be in balls', *berkukes* 'large grained couscous' … PM ʙᴇᴀᴜssɪᴇʀ *brks* 'roll the *brkūks*'.[27]

123. VRṬS/Ṣ *vaṛṭas/vaṛṭaṣ* 1. 'be very closely shaved'. 2. 'completely shave someone's head'. 3. 'be without horns'. *ävəṛṭas* 'without horns (of an animal which should have them)'. Berb ɴᴀɪᴛ-ᴢᴇʀʀᴀᴅ FRḌS(1): ᴅᴀʟʟᴇᴛ *fferḍes* 'have ringworm; be de-horned (ox, ram, billy goat)', *aferḍas* 'ringworm; ringwormed' … PM ʙᴏʀɪs *foṛṭâs* 'ringwormed'. ʙᴇᴀᴜssɪᴇʀ 'ringwormed; have both horns removed (ox)'.[28]

124. ŠNTV *šäntäv* 'remove piece by piece'. *äšäntûv* 'dry hair, not buttered'. ▲Berb ɴᴀɪᴛ-ᴢᴇʀʀᴀᴅ CNTF(1): ᴅᴀʟʟᴇᴛ *acentuf* 'neglected hair, in disarray' … PM ᴄᴏʟɪɴ *šəntəf* 'scratch someone leaving shreds of skin hanging; rip; shred; fluff one's feathers', *šəntūf* 'toupee, topknot of hair'; ʙᴇᴀᴜssɪᴇʀ *šntf* 'form a hoopoe (feathers)'.

125. ŠBŠB *šäbšäb* 'dishevel'. *äšäbšūb* 'tuft of hair on a camel's hump'. ▲ ᴅᴀʟʟᴇᴛ *acebbub* 'hair', *acebcub* 'tuft'; 'crest of feathers'. ʙᴇᴀᴜssɪᴇʀ *šbšŭbt* 'plant plume'.[29]

126. KNZẒ *känẓaẓ* (var. S.-W. of *käntaṭ*) 'bite hard'. *äkänẓâẓ* 'rope under the muzzle'. ▲MA ᴄᴏʟɪɴ *kənẓəẓ* 'shrivel with cooking; become stunted (baby)'.

127. ḤRṬN *ḥaṛṭan* 'become mixed (thoroughbred horses)'. *ḥaṛṭâni* 'freed slave'. ▲Zen *äḥaṛḍan* 'freed slave'. ᴄᴏʟɪɴ: MA *ḥaṛṭāni* to be compared to Berb *āḥaṛḍān* 'black slave'.[30]

128. VGRŠ *tvägräš* 'have shown oneself to be active; be brave'. *ävūgrâš* 'brave, virile'. ▲Zen *äbɣaš* 'courageous man'. MA Essaouira (Moscoso 2002): *fŭgrəš* 'grow, develop (baby)', *fŭgrūš* 'child between 2 and 14 years old'.

129. DRMZ *därmäz* 'be completely sheared'. *ädärmâz* ꜰ *tädärmâzət* 'which has no horns (sheep, goats)'. ▲MA ʟᴏᴜʙɪɢɴᴀᴄ *darmāž* same meaning.

130. ZGLM *zägläm* 'roll (thunder)'. masd. *äzäglâm* – (rare) var. of *tzäglîm*. ▲PM ʙᴇᴀᴜssɪᴇʀ, ᴄᴏʟɪɴ *zəgləm* 'roll'.

4.2.2   In Classical Arabic

In contrast to correspondences with Maghreb Arabic, those with Classical Arabic are not as clearcut. Furthermore, the question mark is also justified, in the last three cases, by the existence in Zenaga of forms with a strong Berber identity (note the *ẓ* in two cases).

---

27. For Madouni, *bərkəš* 'roll the couscous, the *bərkūkəs* (semolina rolled in large grains)' is a borrowing from Berber.

28. The meaning of *farṭasa* in Classical Arabic is quite different.

29. A quite different meaning was noted by Holes in the Orient.

30. On *ḥaṛṭan, ḥṛāṭin* and *āḥaṛḍān*, see Taine-Cheikh (1989: pp. 95–96).

131. MWNK *ṃownäk* 'install comfortably'; *tṃownäk* 'comfortably install one-self'; masd. *tämānûkəṭ*. ▲Zen *yäṃṃunäg* 'live in luxury'. *tmänəkt* 'opulence, wealth'. Cf.? cl. ʔNQ *ʾaniqa* 'admire; like, find something good and nice' (see Basset 1909: p. 18).

132. QWVV/ƔWVV *qawväv/ɣawväv* 'have lots of hair (humans)'. *tāqāwâvət/ tāɣāwâvət* 'hair which is too long'. ▲ Cf.? Cl. *qūfa* 'hair on the back of the head which covers the nape of the neck' and *qaffa* (one of the meanings) 'stand on end (hair, from fear)'.[31]

133. KLKL *kälkäl* 'put a poultice'. *tākəlkəlt/tākəlkâlət* 'type of poultice, medical preparation based on butter and seeds'. ▲Zen **takdyékal* same meaning (Leriche 1953). Cf.? Cl. *kalkala* 'swell, fill with air'.

134. GRMṢ *gaṛmaṣ* 'pinch (with the fingernails)'. ▲Zen *yäškaṛmaẓ* 'pinch'. Cf.? Cl. QRṢ *qaraṣa* same meaning.

135. TWRG *towräg* 'prevent (animal) from fully quenching its thirst'. *tūrgît* PL *tūrgâtən* 'collarbone'. ▲Zen *tūrgiʔđ* and Ouargli *tragla/tragda* 'collarbone'. Cf.? Cl. RQW *tarquwa* 'collarbone'.

136. ẒẒMY *ẓaẓmä* 'have asthma, have an asthma attack', masd. *tāẓəẓmît*. ▲Zen *tänuẓẓəmt* 'asthma'.[32] Cf.? Cl. *zažma* 'labored breathing of a woman giving birth'.

4.2.3   *In other languages*

137. KWTY *mətkowti* (rare) 'have the strength of a monitor lizard'. *kûti* (without article) 'monitor lizard'. ▲Zen *kuđih* same meaning. Berb Tach DESTAING *äkä* PL *äkāten* 'very large poisonous lizard'. Cf.? Wolof *nkoti gā* 'crab' (Basset 1909: p. 229).

138. MYLẒ *mäylaẓ* 'serve as an oral interpreter (*ämäylâẓ*)' + 139. MWLẒ (colonial vocabulary) *mowlaẓ* 'serve as an interpreter (*ämälâẓ*)'. ▲Tu FOUCAULD *îles* 'language; person speaking for, (by extension) interpreter'. Cf. Hebrew *mēlīṣ* same meaning (Vycichl 2005: p. 5).

---

31. In Ḥassāniyya one also finds *gəffä* 'hair which is abundant, long', but this does not explain the presence of *q/ɣ* in the quadriliteral root.

32. This maṣdar could be derived from the Berber root ẒM (*yuẓmä* 'press').

4.3   Complex etymologies

4.3.1   *A note on derivatives in* sa-
In Ḥassāniyya there are quite a few derived forms bearing the prefix *sta-*. Some, with reflexive meaning, are associated with factitive forms in *sa-*. The emergence of these forms in *sa-* was certainly favored by the existence of the Berber factitive derivation in '*s*', but it is used with both Arabic (cf. GBL) and Berber (cf. GWY) roots:

139. GBL *gəblä* 'one of the points of the compass (south or west depending on the region)': *sägbäl* 'place in direction of the *gəblä*'; *stägbäl* 'go in direction of the *gəblä*'.
140. GWY *īggîw* 'griot, musician' (Zen *īggiwi* PL *āggūn* 'griot'; To PRASSE ᵂᵞ*aggu* 'griot') 'turn into a griot (tr. and intr.); *stägwä* 'play the griot'.

However the roots of these verbs in *sa-* are triliteral and therefore only indirectly concern us here (for more details, see Taine-Cheikh 2003).

4.3.2   *Arabic root and Berber formant*
The presence of a Berber formant in *šäktäb* seems quite obvious ('*s*' is usually pronounced *š* in Zenaga). This is also plausible for the other verbs: the realizations *ṣ* and *z* are probably due to the emphatic *ṭ* or the voiced *z* in the root.[33]

141. ŠKTB *šäktäb* 1. 'hone, sharpen (e.g. a pencil)'. ▲ Cl. *kataba* 'write'. For the formant: Zen *äššäktub* 'pencil'; FOUCAULD *sekteb* 'make write'.
142. ŠKTM *šäktäm* (very rare) 'repress (by word, by act)'. ▲ Cl. *katama* 'hide'.
143. ŠLWD *šälwäd* 'swing (child, sling …)'. ▲ Hass. *läwwäd* 'seek', but BORIS *lauwad* 'swing someone around s.t., make go around s.t.'.
144. ṢLBṬ *ṣalbaṭ* 1. 'play *ṣalbaṭ*, which resembles the game knucklebones'. ▲ Cl. *labaṭa* 'throw s.o. on the ground'.
145. ZNVX *zänvax* 'swell due to an illness, bite, anger'. ▲ Cl. *nafaxa* (one of the meanings) 'swell, pump up (arm muscles)'. Hass *nvax* 'blow'.

4.3.3   *Complex cases of (re)borrowings*
Above we saw cases where (i) a Ḥassāniyya denominative verb was (re)borrowed from Zenaga, with the noun being originally Berber; (ii) a noun of Arabic origins used in Ḥassāniyya in a Berberized form, suggesting a borrowing from Zenaga. The following two cases show particularly complex cross interferences.

---

33. Marçais gives several examples of the same type in Djidjelli Arabic, e.g. *serwa* 'flood' as compared to *rwa* 'wet'.

146. SƔNN. This root is a special case.[34] On one hand, the quadriliteral verb *säɣnän* 'mix gum with water (to make ink)' appears to be calqued on the Zenaga *yässuɣnän* 'thicken (ink) with gum' – the factitive form of *yuɣnän* 'thicken (when gum is added, for ink)'.

On the other hand, the Zenaga noun *əssayan* 'gum (medicine)', despite possible relations to the Ouargli *tayənyant* 'type of pungent resin used as incense', appears to be a borrowing from Arabic, with assimilation of the definite article (whence the geminate *s*). There has probably been interference from the Ḥassāniyya *sämɣa* 'ink (local)' (< Cl. *ṣamɣ* 'Arabic gum'), but the assimilation is not complete (cf. *ɣ + n/m + ɣ*).

The divergence between the root of the verb and that of the noun, in both Ḥassāniyya and Zenaga, is the consequence of this double interference.

147. MWŽR. Berber origins for the quadriliteral verb *ṃowžär* 'have for tributary; ransom' are probable, given, on one hand, the root MZR in Tuareg (FOUCAULD *mezeṛ* 'defend'), and on the other the correspondence between Hass *äṃäžâr* 'protector' and Zen *ämäz̄är* 'Emir, lord, protector (of the people)'. The existence in Zenaga of the verb *yäz̄är* 'protect' is in keeping with this hypothesis, but the fact that the Tuareg and Zenaga verb forms do not coincide (and their isolation in Berber) is troubling. Lastly one may posit that this quadriliteral root is more closely tied than would at first appear to the Arabic root ŽWR (*žär(a)* 'be unfair to s.o.'), especially as the notion of protection is closely linked in Moorish society to extortion, as evidenced by the translation of *ṃowžär*.

5.   Conclusion

Of all the quadriliteral verbs found in Ḥassāniyya, I studied the roots sharing at least one trait with Berber (a linguistic feature – phonetic and/or morphosyntactic – or a lexical token), i.e. approximately one fifth of the quadriliteral roots.[35] Excluding the two triliteral roots (#139 and 140), this leaves 146 roots. There are several different types of interference with Berber contributing only an affix (#141 to 145), or giving tit for tat (#146 and 147). In total, roots of Berber origin represent between 7.3% and 18.25% of all quadriliterals (the lower figure being based on the first 59

---

34. Its complexity is comparable however to that of the Berber root NR (Tu FOUCAULD *ener* 'guide') which gave namely 'guide (v. and n.)', 'find one's direction' in Ḥassāniyya and, in Zenaga, the surprising doublet *yinär/yišnär* 'find one's direction' (Taine-Cheikh 2003: p. 115).

35. This is much lower than the set of quadriliteral verbs of unknown origin because there was no reason to suspect interference from Berber in the case of verbs such as *bäwžäž* 'fumigate' or *däwšäl* 'knock unconscious'.

roots) – percentages which in any case are quite low in an area where, given the lack of frequent correspondences with Classical Arabic or other Arabic dialects, one would expect much higher figures.

In most cases, the Ḥassāniyya quadriliteral verbs appear to have been formed on nominal borrowings, but the target form may have been verbal in some cases, namely in 34 out of the first 59 cases – which is considerable given that nouns are the most easily borrowed.

Whatever the percentage retained, these are "cultural borrowings" rather than "core borrowings" (Myers-Scotton 2002: p. 239). Two semantic fields dominate: animals (husbandry, riding, doctoring), and illnesses. Together these two fields represent 65 of the quadriliteral roots studied here (and 28 of the first 59). Out of the 62 other verbs representing actions, slightly less than half are habitual actions of a traditional type (e.g. #7 'far away hunt' or 14 'beat down (gum)'), and slightly more than half are more general actions (e.g. #10 'lock up', #18 'go on foot'). The 19 other quadriliterals belong to more or less specific semantic fields such as physical traits (e.g. #2 'odd-eyed'), social features (e.g. #38 'mistreat a slave') and relations to time (e.g. #53 '(make) leave the afternoon'). Concerning ties between Arabic and Berber, all cases are represented: the Berber lexical meaning is often retained or is identical, but there can also be significant shifts in meaning. When there is divergence, the target language often shows specialization, especially when the borrowing is a verb.

As for the formation of quadriliterals one notes – with the exception of some unforeseeable evolutions (cf. 2.2.2) – a tendency to retain derivational affixes (cf. 2.3), at times even integrating them in triliteral roots of Arabic origins (cf. 4.3). Among the surprising features of these quadriliterals are the absence of reduplicated biliterals and the high frequency of one or even two root glides. In fact it is often by addition of a root W/Y that a Berber bi- or triliteral root becomes a quadriliteral Ḥassāniyya root, not only for borrowed nominals but also borrowed verbs (e.g. #7 GYMR < GMR 'far away hunt').

On the whole, what is striking is the complexity of relations between Zenaga and Ḥassāniyya, with lexemes which are often specific to this area, making it difficult to do more than hypothesize various borrowings and borrowing processes. The present study does show however that there are lexemes which are found only in the Sahara area (e.g. #12 'fennec') or in Maghreb Arabic (see in particular 4.2.1) and the retention in Ḥassāniyya of various terms of Berber origin not (or no longer) found in Zenaga.

References

Akkuş, F. (2016). The Arabic Dialect of Mutki-Sason Areas. In G. Grigore & G. Biţună (eds.), *The Arabic Varieties: Far and Wide. Proceedings of the 11th International Conference of AIDA – Bucharest* (2015). Bucarest: Editura Universităţii din Bucureşti. 29–40.

Albader, Y. B. (2016). Quadriliteral Verbs in Kuwaiti Arabic. In G. Grigore & G. Biţună (eds.), *The Arabic Varieties: Far and Wide. Proceedings of the 11th International Conference of AIDA – Bucharest* (2015). Bucarest: Editura Universităţii din Bucureşti. 53–63.

Basset, R. (1909). *Étude sur le dialecte zénaga. Mission au Sénégal*. Paris: Leroux.

Beaussier, M. (1958). *Dictionnaire pratique arabe-français* Alger: La Maison des Livres.

Behnstedt, P. and M. Woidich (2011). *Wortatlas der arabischen Dialekte (WAD). Band 1. Mensch, Natur, Faune, Flora*. Leiden: Brill.

Boris, G. (1958). *Lexique du parler arabe des Marazig*. Paris: Klincksieck.

Colin, G. S. (1926). Notes de dialectologie arabe. *Etymologies maġribines. Hespéris* 6, 55–82.

Colin, G. S. (1930). Notes de dialectologie arabe. *II. Sur l'arabe marocain de l'époque almohade. Hesperis* 11, 104–120.

Colin, G. S. (1993 [1993–1997]). *Le Dictionnaire Colin d'Arabe Dialectal Marocain*, édité par Zakia Iraqui-Sinaceur. Rabat–Paris: Al Manahil.

Dallet, J.-M. (1982). *Dictionnaire kabyle-français, parler des At Mangellat, Algérie*. Paris: SELAF.

Delheure, J. (1987). *Dictionnaire ouargli–français*. Paris: SELAF.

Destaing, E. (1920). *Etude sur la Tachelḥît du Soûs. I Vocabulaire français-berbère*. Paris: Leroux.

Destaing, E. (2007 [1914]). *Dictionnaire français–berbère (Dialecte berbère des Beni-Snous)*. Paris: L'Harmattan.

El Aissati, A. (2006). Berber Loanwords. In K. Versteegh (editor-in-chief), *Encyclopedia of Arabic Language and Linguistics*. Leiden: Brill. 293–299.

Foucauld, C. de (1951–52). *Dictionnaire touareg-français (Ahaggar)*. Paris: Imprimerie Nationale de France.

Guignard, M. [2005 (1ère éd. 1975)]. *Musique, honneur et plaisir au Sahara*. Paris: Geuthner.

Heath, J. (2004). *Hassaniya Arabic (Mali)–English–French Dictionary*. Wiesbaden: Harrassowitz.

Heath, J. (2006). *Dictionnaire touareg du Mali. Tamachek–anglais–français*. Paris: Karthala.

Holes, C. (2004). Quadriliteral verbs in the Arabic dialects of eastern Arabia. In M. Haak, R. de Jong & K. Versteegh (eds.), *Approaches to Arabic dialects: Collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday* Leiden–Boston: Brill. 97–116.

Kamil, M. (1963). *Beiträge zur Entstehung der vierradikaligen Verben in den gesprochenen semitischen Sprachen*. Mémoires de l'Institut d'Égypte, 57. Le Caire.

Kazimirski, A. de Biberstein (1860). *Dictionnaire arabe-français* 2 tomes. Paris: Maisonneuve et Cie.

Kossmann, M. (2013). *The Arabic Influence on the Northern Berber Languages*. Leiden: Brill.

Lanfry, J. (1973). *Ghadamès II. Glossaire.* Alger: Le Fichier Périodique.

Laoust, E. (1920). *Mots et choses berbères*. Paris: Challamel.

Lentin, J. (2010). Sur quelques préformantes utilisées dans la morphogénèse de la racine: l'exemple de l'arabe. In F. M. Fales & G. F. Grassi (éds), *Camsemud 2007. Proceedings of the 13th Italian meeting of Afro-asiatic linguistics held in Udine*, May 21st–24th, (2007). 281–295. Padova: S.A.R.G.O.N. Editrice e Libreria.

Leriche, A. (1953). Phytothérapie maure. *Mélanges Ethnologiques*. Dakar: IFAN. 265–306.

Loubignac, V. (1952). *Textes arabes des Zaër. Transcription, traduction, notes et lexique*. Paris: Librairie orientale et américaine Max Besson.

Madouni, J. (2001). À propos de verbes quadrilitères dans un parler de l'Ouest algérien (Sidi-Bel-Abbès). *Estudios de Dialectología Norteafricana y Andalusí* 5, 243–251.

Marçais, P. (1956). *Le parler arabe de Djidjelli (Nord Constantinois, Algérie)*. Paris: Librairie Adrien-Maisonneuve.

Monteil, V. (1952). *Essai sur le chameau au Sahara occidental*. Saint-Louis: Centre IFAN Mauritanie–Saint-Louis.

Moscoso, F. (2002). El dialecto árabe de Essauira a partir de los textos publicados en 1893 por Albert Socin (parte 1). *Estudios de Dialectología Norteafricana y Andalusí* 6, 35–97.

Myers-Scotton, C. (2002). *Language contact: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.

Naït-Zerrad, K. (1998–2002). *Dictionnaire des racines berbères (formes attestées)* 3 volumes parus. Paris–Louvain: Peeters.

Nicolas, F. (1953). *La langue berbère de Mauritanie*. IFAN–Dakar.

Pellat, C. (1950). Sur deux emprunts au berbère par l'arabe dialectal nord-africain. *Mélanges William Marçais*. Paris: Maisonneuve et Cie. 277–288

Prasse, K. G., A. Ghoubeïd and M. Ghabdouane (2003). *Dictionnaire Touareg–Français (Niger)*. Copenhagen: Museum Tusculanum Press – Université de Copenhague.

Procházka, S. (1993). Some remarks on the semantic function of the reduplicated verbs quadri-literal verb (structure *faˤfaˤa*). In K. Devenyi et al. (eds), *Proceedings or the Colloquium on Arabic Lexicology and Lexicography*. Budapest: Eötvös Loránd University. 97–103.

Taifi, M. (1992). *Dictionnaire tamazight–français (Parlers du Maroc central)*. Paris: L'Harmattan-Awal.

Taine-Cheikh, C. (1988–1998). *Dictionnaire Hassaniyya–Français* 8 volumes parus. Paris: Geuthner.

Taine-Cheikh, C. (1989). La Mauritanie en noir et blanc. Petite promenade linguistique en *ḥassāniyya*. *Revue du Monde Musulman et de la Méditerranée* 54 (1989/4), 90–105.

Taine-Cheikh, C. (1994). Le *ḥassāniyya* de Mauritanie, un dialecte non-marginal de la péri-phérie. In J. Aguadé, F. Corriente and M. Marugán (eds.), *Actas del Congreso Internacional sobre Interferencias Lingüisticas Arabo-Romances y Paralelos Extra-Iberos Madrid* (10–14 déc. 1990). Zaragoza. 173–199.

Taine-Cheikh, C. (1997). Les emprunts au berbère zénaga. Un sous-système vocalique du *ḥassāniyya*. *Matériaux arabes et sudarabiques* 8 (N.S.), 93–142.

Taine-Cheikh, C. (2003). Les valeurs du préfixe s- en Ḥassāniyya et les conditions de sa gramma-ticalisation. In I. Ferrando and J. J. Sánchez Sandoval (ed.), *AIDA 5th Conference Proceedings, Cádiz september 2002*. Cádiz: Servicio de Publicationes Universidad de Cádiz. 103–118.

Taine-Cheikh, C. (2007). Périphérie géographique et perméabilité aux contacts. Le cas du Maghreb. *Romano-Arabica* 6–7, 159–178.

Taine-Cheikh, C. (2008a). Arabe(s) et berbère en contact: le cas mauritanien. In M. Lafkioui & V. Brugnatelli (eds.), *Berber in Contact. Linguistic and Sociolinguistic Perspectives*. Köln: Köppe. 113–138.

Taine-Cheikh, C. (2008b). Les verbes dérivés à préfixe 'S' en berbère: le cas du zénaga, entre grammaire et lexique. In G. Takács (ed.), *Semito-Hamitic Festschrift for A.B. Dolgopolsky and H. Jungraithmayr*. Berlin: Dietrich Reimer. 284–309.

Taine-Cheikh, C. (2008c). *Dictionnaire zénaga–français. Le berbère de Mauritanie par racines dans une perspective comparative*. Köln: Köppe.

Taine-Cheikh, C. (2015). La ronde des saisons … et des labiales ! *Quaderni di Studi Berberi e Libico-Berberi (Studi Africanistici). La lingua nella vita e la vita della lingua. Itinerari e percorsi degli studi berberi. Miscellanea per il Centenario di studi berberi a "L'Orientale" di Napoli. Scritti in onore di Francesco Beguinot* 4, 297–319.

Taine-Cheikh, C. (forthcoming). Onomatopées et iconicité du langage: les racines quadrilitères à redoublement (C1C2C1C2) en arabe maghrébin. In S. El Ayachi (ed.), *Hommage à Miloud Taifi: "Pannes des sens et stratégies de dépannage sémantique"*. Fès, Maroc: Faculté des Lettres et des Sciences Humaines de Saïs-Fès.

Tauzin, A. (1993). *Contes arabes de Mauritanie*. Paris: Karthala.

Vycichl, W. (2005). *Berberstudien & A Sketch of Siwi Berber (Egypt)*. Köln: Köppe.

Wichmann, S. and J. Wohlgemuth (2008). Loan verbs in a typological perspective. In T. Stolz, D. Bakker and R. Salas Paloma (eds.), *Aspects of Language Contact. New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*. Berlin–New York: de Gruyter. 89–121.

Loan verbs in Egyptian Arabic

Perspectives and evidence from social media

Ashraf F. Hassan

University of Bayreuth, University of Naples "L'Orientale"

This paper will explain the strategies of loan verbs integration in Egyptian Arabic (EA). As a recipient language, EA adopts two strategies: (a) insertion with 'Light Verb Strategy'; and (b) Direct Insertion either with or without 'Reduction to Root.'

While direct insertion strategy without 'reduction to root' is used almost exclusively for imperative loan verbs, the same strategy with 'reduction to root' is open to any 'input form'. To each loan verb EA assigns a root and the loan verb assumes one of the EA verbal forms.

An investigation of new loan verbs passed to EA through Social Media, while they are being integrated, gives us further insight, and therefore a better understanding, into the integration process of loan verbs in general.

Keywords: Social Media, Egyptian Arabic, lexical borrowing, loan verbs, integration

1. Introduction

Apart from Mifsud's *Loan verbs in Maltese* (1995), no extensive study has been made of loan verbs in any Arabic varieties. However, a few works have covered this subject, albeit partially, for example Versteegh's *Loan Verbs in Arabic and the DO-construction* (2009) and an article by Hafez (1996) on *Phonological and morphological integration of loanwords into Egyptian Arabic*.

This paper tries to contribute to the understanding of how loan verbs are integrated into EA through an analysis of the language used in Egyptian Social Media, especially Facebook and Twitter, being the most used public interaction platforms in Egypt. Following the statement of Poplack & Sankoff (1984: p. 125) that *"The assimilation of loanwords is, of course, a diachronic process, best studied if possible at several points in time"*, such analysis is based on a long and close observation of

that language, in order to notice the mechanism of verbal borrowing and accommodation in EA.

Facebook first appeared in Egypt February 2004 with an English interface; after five years an Arabic Beta Interface was provided. Twitter, instead, was launched in July 2006 and by March 2012 had an Arabic Interface. The first users of both Social Media forms needed a certain level of competence in English. They wrote predominantly in English, but sometimes employed Arabic with Roman Script. This situation caused Egyptian users to be strongly exposed to the English language, creating a context of language contact between English and EA, which resulted in numerous cases of lexical borrowings.

If we look at the statistics, among about 94 million Egyptians, 35 million have access to the Internet, with 37% penetration.[1] By the early 2017, the number of Facebook active accounts in Egypt amounted to almost 35 million as well, constituting 23% of Facebook users in the Arab region (Salem 2017: p. 35).[2] Over 90% of Egyptian Facebook users posted in Arabic, about 24% in English and less than 2.5% in French. During the last three years, Arabic language gained more ground at the expense of both English and French, where the percentages were 75%, 34% and 4% respectively.[3]

As for Twitter, we have about 1.7 million users in Egypt, constituting 18% of the overall Twitter users in the Arab region and generating 20% of all tweets in the region. The tweets are mainly in Arabic (75%) and English (14%) (Salem 2017: p. 45, 53).

2.   Theoretical background

This article finds its theoretical basis in the work of Jan Wohlgemuth (2009) *A Typology of Verbal Borrowings*. As defined by Wohlgemuth (2009: p. 67), a loan verb is *"an established borrowed lexical item (i.e. not one inserted ad-hoc) which can count as a verb (or is predominantly "verby", i.e. an action word that prototypically serves as the head of a predicate phrase), both in the recipient (borrowing) and in the donor (source) language"*. The language from which a loanword has been borrowed is called the *donor* language, and the language into which it has been borrowed is

---

1.   Source: the report of 'We Are Social' agency *Digital in 2017: Northern Africa*, available online https://www.slideshare.net/wearesocialsg/digital-in-2017-northern-africa

2.   More than 34% of users own more than one Facebook account (Salem 2017: p. 10).

3.   Source: http://www.arabsocialmediareport.com/UserManagement/PDF/ASMR6_En_Final.pdf

the *recipient* language (Haspelmath 2009: p. 36; Wohlgemuth 2009: p. 51). In our case these are English and EA respectively.

Since the source words of loanwords often have phonological, orthographic, morphological and syntactic properties in the donor language that do not fit into the system of the recipient language, loanwords often undergo changes to make them fit better into the recipient language. These changes are generally called loanword adaptation (Haspelmath 2009: p. 42).

Wohlgemuth (2009: p. 293) identifies four main type classes of loan verb accommodation mechanisms, called strategies:

> **Direct Insertion** (DI), where the borrowed verb stem is simply used like a native one without any morphosyntactic adaptation;
> **Indirect Insertion** (IndI), where a verbalizer of some kind is applied so that the loan verb can then be inflected.
> **Light Verb Strategy** (LVS), where a borrowed verb is to enter it as a non-inflecting part into a complex predicate, joining a native verb which takes all the inflection.
> **Paradigm Insertion** (PI), where the borrowed verb's inflectional morphology of the donor language is borrowed along with it, introducing a new inflectional paradigm into the recipient language.

Additionally, Wohlgemuth (2009: p. 178) states that in Arabic: *"The borrowed verbs must normally be transformed to a root of three (occasionally four, rarely two or five) consonants. These roots can be combined with different inflectional and derivational templates to produce verbs, nouns, adjectives and their inflected forms. Many of these roots and their basic citation forms already have "verby" semantics. Further (formal) verbalizing derivation is thus not necessary […] This accommodation technique is subsumed under Direct Insertion".*[4]

3.   Social Media and loan verbs

As mentioned above, Facebook and Twitter have been used for some years with the English interface before having an Arabic one. This situation required some competence in English among the old users, or a "reasonably widespread bilingualism", which explains the widespread use of loanwords for new concepts (See Haspelmath 2009: p. 47).

Social Media, contributed in the borrowing of verbs in two directions:

---

4. For more discussion on the particularity of Semitic languages, see Wohlgemuth (2009: pp. 92–3 and pp. 173–8).

1. the extension of old loan verbs: (a) just as they were used before: /ji-ʕmil ʔab-deet/ 'to update',[5] /ji-haak, ji-thaak/ 'to hack, be hacked'; or (b) extending their use to more semantic domains: /ji-hannig/ from 'to hang, to freeze (in computing), to stop working (in general)';

2. the introduction of new loan verbs', either (a) relating to Social Media interfaces: /ji-ʃajjir/ (share); (b) relating to other domains: /ji-kraaʃ/ 'to have a crush on someone).

4.   Integration strategies

Generally speaking, the integration of loan verbs in EA adopts two strategies: (1) the Light Verb Strategy (LVS), or (2) the Direct Insertion a. without 'Reduction to Root' (DI), or b. with 'Reduction to Root' (DIRR).

4.1   Light Verb Strategy (LVS)

Many loan verbs enter EA as a non-inflecting part into a complex predicate, joining a native verb which takes all the inflection. The preferred native verb is /ʕamal, ji-ʕmil/ 'to do, make' (1) and, less common, /xad, ja-axud/ 'to take' (2). Sometimes, they alternate for the same loan verb (3):

(1) /ji-ʕmil fevorit/ 'to favorite';[6] /ji-ʕmil folo(baak), ɑnfolo/ 'to follow (back), unfollow'; /ji-ʕmil mijuut/ 'to mute'; /ji-ʕmil ṣabtiwiit/ 'to subtweet'; /ji-ʕmil trend/ 'to trend'; /ji-ʕmil ɑbdeet/ 'to update'; /ji-ʕmil ɑnfrend/ 'to unfriend'; /ji-ʕmil ʔaad/ 'to add'; /ji-ʕmil blokk/ 'to block'; /ji-ʕmil dawinlood/ 'to download'; /ji-ʕmil ribort/ 'to report (i.e. another user)'; /ji-ʕmil pook / 'to poke'

(2) /ja-axud kootritwiit/ 'to quote retweet'; /ja-axud skriinʃott, snapʃott/ 'to screenshot, snapshot'[7]

(3) /ji-ʕmil, ja-axud kobi-w-best/ 'to copy & paste'; /ji-ʕmil, ja-axud kootritwiit/ 'to quote retweet'

---

5. For the phonemic transcription between the two slashes, I use the International Phonetic Alphabet (IPA) in my phonemic transcription, followed by the translation into English, which is the source word for the loanword.

6. Recently, "favorite" has been replaced by "like" in Twitter Interface.

7. One may argue that the two are examples of calque from English "take screenshot, snapshot".

We have other light verbs that are used in particular contexts, such as /katab, ji-ktib/ 'to write' (4), /baʕat, ji-bʕat/ 'to send' (5), /xabaṭ, ji-xbaṭ/ 'to hit, click' or /daas, ji-duus/ 'to hit, click' (6). In most cases, these verbs can, also, be substituted with /ʕamal, ji-ʕmil/:

(4)   /ji-ktib (ji-ʕmil) komment/ 'to comment'; /ji-ktib (ji-ʕmil) riblaaj/ 'to reply'

(5)   /ji-bʕat di(i)-emm/ 'to direct message';[8] /ji-bʕat (ji-ʕmil) menʃan/ 'to mention'

(6)   /ji-xbaṭ, ji-duus (ji-ʕmil) lajik/ 'to like'; /ji-xbaṭ, ji-duus (ji-ʕmil) fevorit/ 'to favorite'

4.2   Direct Insertion without 'Reduction to Root' (DI)

The Direct Insertion strategy (DI) without 'Reduction to Root' is commonly used where the 'input form' in the donor language is an imperative verb as in (7):

(7)   /lajik/ 'like!'; /ʃeer/ 'share!'; /ritwiit/ 'retweet!'; /folo/ 'follow!'

This strategy is also attested in old loan verbs in EA that originate from the 19th century or the beginning of 20th century, especially those related to the maritime language. We find, for instance, /majna/ < from it. ammaina > 'haul down!' and /ʔitraaka/ < from it. attracca > 'moor!' (Spiro 1904: p. 21, 24).

Yet, in the last decade, due to the influence of Social Media, we witness a new trend gaining popularity. This refers to a new wave of lexical borrowing, which is not easy to define, since the phenomenon is in its early stage. Many verbs are being borrowed through the DI strategy but their 'input form' is not only the imperative verb (8):

(8)   /ji-krob (krop)/ 'to crop (an image)'; /ji-ʔanfolo/ 'to unfollow'; /ji-twist/ 'to twist'; /ji-twiit/ 'to tweet'; /ji-ritwiit/ 'to retweet'; /ji-troll/ 'to troll'; /ji-trend/ 'to trend'; /ji-ṣabtwiit/ 'to subtweet'; /ji-skajib/ 'to skype'; /ji-serʃ/ 'to search'; /ji-flert/ 'to flirt'; /ji-ʔorgazm/ 'to orgasm'; /ji- dawinlood/ 'to download'; /ji-stolk/ 'to stalk'

This innovative phenomenon in EA, probably, originated with a famous slogan of Pepsi-Cola in Egypt which was introduced in 1996:

(9)   /jalla   ʃabaab   jalla   n-pepsi/
      let's (go) guy.PL   let's   1PL-IPFV.pepsi
      'Let's go guys! Let's pepsi!'

---

8.   We may find the acronym /di(i)-emm/ in Arabic or Roman script as a calque of 'DM'. The same happens as well with /loll/ 'LOL, Laugh(ing) Out Loud'.

These verbal roots are not productive (i.e. they have no verbal nouns, participles … etc.) and they exhibit partial morphological integration. They conjugate only in imperfective/infinitive form (10) and, as far as I know, no perfective conjugation such as /*troll-it/ 'she trolled' is found. The imperfective prefixes /bi-/ and /ha-, ħa-/ and the direct object suffixes can be attached to them (11).

(10)   /lamma ʔa-flirt/ 'when I flirt'; /jareet ni-legalajiz/ 'wish we can legalize'; /enta bi-t-ṣabtwiit ʕala miin/ 'whom are you (m.) subtweeting for?'; /ti-blokk-i lli nti ʕajzaa/ 'you (f.) block whoever you want'; /ji-lebol in-naas/ 'he labels the people'; /bi-t-troll nafs-a-ha/ 'she trolls herself'; /ʕaʃaan ji-ʔimpres-u l-ʔalmaan/ 'so that they impress the Germans'

(11)   /ha-a-meks-u b-raaħt-i/ 'I will mix it on my own'; /ħa-n-ritwiit biʃedda; 'we will retweet strongly' /bi-ji-ristɑrt ir-rɑwtɑr/ 'he is restarting the router'

In fact, we deal here with ambiguous instances. On the one hand, it is difficult to consider them as code-switches, since the violate the 'free-morpheme constraint' formulated by Poplack (1980: p. 585): "*codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme.*" Not only, but these instances are, to various extents, widespread among the Social Media users, even those who are monolingual.

On the other hand, these instances are not completely integrated, at least morphologically, so we cannot define them as "established" borrowings. Perhaps, that is the case of what Haspelmath (2009: p. 41) called 'incipient loanwords', and it is a matter of time to get integrated completely if accepted by the other members of the community of speakers: "*It is in fact reasonable to assume that as a borrowed word is more and more used, it tends to become phonologically and morphologically integrated, to displace competing recipient language forms, and at least eventually, to be accepted by its native speakers*" (Poplack & Sankoff 1984: p. 105).

This new trend might be an intermediate step before integrating these loan verbs through the DIRR pattern (see below) or an introduction of a new pattern in EA, due the increasing impact of English on EA. This impact is also demonstrated in the introduction of the plural morpheme in English *-s/-es* in many nominal loanwords: /admin, adminz/ 'admin(s)'; /pedʒ, pedʒiz/ '[internet] page(s)'; /faanz/ 'fans'; /kanz/ '[beverage] cans'.[9]

This new trend goes side by side with an analogue trend that is also spreading amongst EA speakers, i.e. the neologizing of denominal verbs which derive from locutions and do not adhere fully to the verbal system of EA. I give here two examples: /ji-kossomm/ 'to *motherfucker* s.o.', derived from the locution

---

9.   It is integrated as a singular noun in EA.

/koss/ + /umm/ + GEN 'mother's vagina, lit. the vagina of the mother (of s.o.)' and /ji-maʕleʃʃ/ 'to disregard, to pay no attention to s.o., derived from /maʕleʃʃ/ 'never mind, don't worry about it'. These neologisms do not fit any verbal form of EA.[10] Though they are used very often.

4.3   Direct Insertion with 'Reduction to Root' (DIRR)

Wohlgemuth (2009: p. 175) states that, for all varieties of Arabic, "*Direct Insertion is the most widespread strategy, and it is chiefly represented by the pattern type of Reduction to root*". Furthermore, the observation of the process of loan verb accommodation in EA suggests that the DIRR is the ultimate point of accommodation: it is as 'slim' as the loan verbs accommodated through the DI strategy, yet they keep the full morphosyntactic functionality as those of the LVS.

In this accommodation pattern the loan verbs, which usually have been already accommodated through one of the aforementioned strategies, reshape into roots to conform to the EA morphophonological requirements of templatic verb inflection (see Wohlgemuth 2009: p. 92). The verbs which are accommodated through a DIRR pattern assume only a triliteral or quadriliteral root.

Whenever the model verb in English contains only two consonants, the shaping of the root in EA is determined by the vocalic/consonantic quantity in the model verb in English. When the model verb has a long vowel (or a semivowel), that vowel will act as a radical, resulting in *verba mediae infirmae* (12a). And if it has a long consonant (geminate), it will act as two radicals resulting in *verba mediae geminatae* (12b):

(12) a.   /haak, ji-haak/ 'to hack' (in EA /haak/)
      b.   /rall, ji-roll/ 'to roll' (in EA /roll/)

Equally, if the model verb has more than four consonants, only four will be maintained and the excess will be elided (13):

(13)   /rastar, ji-rastar/ 'restart' (with the elision of the second 't')

The process of reshaping the model verb into roots goes along with the process of assuming an EA verbal form. Principally, the loan verbs prefer one EA verbal form for the triliteral roots (14) and one analogue form for the quadriliteral roots (15) with their respective reflexive/passive forms. It is worth observing that, at a phono-morphological level, these two pairs are quantitively identical.

---

10. When degeminated, /ji-maʕleʃʃ/ would fit into a quadriliteral root verb, still it did not.

(14)   /ballik, it-ballik/ 'to block, be blocked'; /dallit, it-dallit/ 'to delete, be deleted'; /ʃajjir, it-ʃajjir/ 'to share, be shared'; /kawwit, it-kawwit/ 'to quote, be quoted'; /sarriʃ, it-sarriʃ/ 'to search, be searched for'; /sajjiv, it-sajjiv/ 'to save, be saved'; /tawwit, it-tawwit/ 'to tweet, be tweeted'

(15)   /gawgil, it-gawgil/ 'to google, be googled'; /haʃtig, it-haʃtig/ 'to hashtag, be hashtaged'; /manʃin, it-manʃin/ 'to mention, be mentioned'; /ratwit, it-ratwit/ 'to retweet, be retweeted'

	Perfetive	*Imperfective*
Triliteral Roots		[v: /a, i/]
Active/Transitive	$C_1aC_2C_2vC_3$	$ji\text{-}C_1aC_2C_2vC_3$
Passive/Reflexive	$it\text{-}C_1aC_2C_2vC_3$	$jit\text{-}C_1aC_2C_2vC_3$
Quadriliteral Roots		[v: /a, i/]
Active/Transitive	$C_1aC_2C_3vC_4$	$ji\text{-}C_1aC_2C_3vC_4$
Passive/Reflexive	$it\text{-}C_1aC_2C_3vC_4$	$jit\text{-}C_1aC_2C_3vC_4$

Likewise, the loan verbs, with the exception of a very few cases (see 13 above), opt also for one of the two possible vowels in the last syllable, i.e. the front close vowel /i/, and not the open one /a/.

However, some verbs such in (12) opted for the basic verbal form,[11] i.e. $C_1vC_2vC_3 - ji\text{-}C_1C_2vC_3$ with its variants $C_1aaC_3 - ji\text{-}C_1vvC_3$ (*verba mediae infirmae*) and $C_1vC_2C_2 - ji\text{-}C_1vC_2C_2$ (*verba mediae geminatae*).

For the case of /haak, ji-haak/ 'to hack', the front open long vowel /aa/ cannot constitute a radical. It is, rather, a mutation of the radical semivowels /w, j/ or their respective long vowels /uu, ii/. These semivowels do not appear in the perfective of the basic verbal form, but they do in the imperfective, the verbal noun, and other verbal forms of the same root.

No semivowel has been assigned to the verb /haak, ji-haak/ 'to hack', as mutated radical, unlike the case of /ʃajjir, ji-ʃajjir/ < from /ʃeer/ > 'to share' where the front close-mid long vowel suggests typically a mutation of the semivowel /j/; or the case of /tawwit, ji-tawwit/ < tiwiit > 'to tweet' where a semivowel /w/ already exists.

The case of /rall, ji-roll/ 'to roll' constitutes an exception, as the approximant /l/ does not follow the trill /r/ in EA roots.[12] That may explain why doubling the approximant /l/ would not sound good.

Another curious case is that of the model verb 'block' which has been borrowed and accommodated into EA through various strategies:

---

11.   See Woidich (2006: pp. 60–61) for more details on this form.

12.   In Classical Arabic, this occurs in few roots: *ʾ-r-l*; *b-r-l*; *g-r-l*; *ġ-r-l*; *w-r-l* (see: al-ʿAskarī 1988: p. 1/396). But no one of them passed into EA lexicon.

a. In an early stage through LVS /ji-ʕmil blokk/, through DI having either imperative as model verb /blokk/, or not /ji-blokk/

b. Later, through DIRR with two variants: triliteral /ballik, ji-ballik/ (root: *b-l-k*) and quadriliteral /balwik, ji-balwik/ (root: *b-l-w-k*).

The triliteral root form is more common than the quadriliteral one. But, why does the latter exist at all? The model form has only three distinguished consonants, no long vowel nor semivowel. So, it should reshape into a triliteral root. We may find the answer in the way this loan verb spread in the early stage: in all the three ways above mentioned, i.e. /ji-ʕmil blokk/, /blokk/, and /ji-blokk/, the loan verb vowel has been transcribed in EA with the grapheme *wāw* <و>. This may have misled users who had less competence in English and they may have thought the *wāw* represented a semivowel, rather than a short vowel.

5.   Conclusive remarks

Since the appearance of various Social Media, loan verbs in EA (as much as other lexical categories) increased significantly. The *donor language* of such loan verbs is predominantly English, as this is the dominant language of Social Media interfaces. Nevertheless, in sporadic cases we have French as a *donor language*, as for instance /ʃapoo/ < from fr. Chapeau! > 'hat, bravo!'.

The loan verbs accommodate in EA as *recipient language*, through two strategies:

1. the insertion with 'Light Verb Strategy'
2. the Direct Insertion
    a. without 'Reduction to Root'
        i.   with an imperative *model verb*
        ii. with non-imperative *model verb*
    b. with 'Reduction to Root'

Diachronically speaking, the insertion with 'Light Verb Strategy' and Direct Insertion without Reduction to Root with an imperative *model verb* appeared first, once it was made possible to write in Arabic on Social Media. Afterwards, Direct Insertion without Reduction to Root with non-imperative model verb followed. However, it does not seem to be firmly established yet. Then, and at ultimate step, came the Direct Insertion with Reduction to Root as the final goal of loan verbs accommodation.

Some loan verbs accommodate through more than one of – or even all – the aforementioned strategies and patterns, like for instance 'to block' (see above) or 'to retweet': /ji-ʕmil ritwiit/, /ritwiit/, /ji-ritwiit/ and /ji-ratwit/. Nonetheless, the

Direct Insertion with Reduction to Root is steadily gaining ground at the expense of other strategies and patterns. Still, this does not mean that the other accommodated forms of the same loan verb will disappear completely. Most of them are still in use although to a lesser degree, and some verbs which took part of older common sayings continue to exist. For instance, the model verb 'to share' is commonly used in the form accommodated through DIRR /ʃajjir, ji-ʃajjir/, but it keeps the form accommodated through LVS /ʃeer/ in the expression /ʃeer fi-l-xeer/ 'Share, for goodness' sake!', because it rhymes with the noun /xeer/.

References

al-ʿAskarī, A. (1988). *Ǧamharat al-amṯāl*, edited by Muḥammad A. Ibrāhīm & ʿAbd al-Maǧīd Qaṭāmiš, 2nd ed., 2 vols. Bayrūt: Dār al-Ǧīl.

Hafez, O. (1996). Phonological and morphological integration of loanwords into Egyptian Arabic. *Égypte / Monde arabe* 27–28: 383–410.

Haspelmath, M. (2009). Lexical borrowing: Concepts and issues. In M. Haspelmath and U. Tadmor (eds.), *Loanwords in the World's Languages: A Comparative Handbook* (pp. 37–54). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110218442.35

Haspelmath, M., and U. Tadmor (eds.) (2009). *Loanwords in the World's Languages: A Comparative Handbook*, Berlin: Mouton de Gruyter.

Hinds, M., and E. Badawi (1986). *A Dictionary of Egyptian Arabic*. Beirut: Librairie du Liban.

Mifsud, M. (1995). *Loan verbs in Maltese: A descriptive and comparative study*. Leiden, Boston: Brill.

Moravcsik, E. (1975). Verb borrowing. *Wiener Linguistische Gazette* 8: 3–30.

Muysken, P. (2008). *Functional Categories*. Cambridge: Cambridge UP. https://doi.org/10.1017/CBO9780511755026

Poplack, S. (1980). Sometimes I'll start a sentence in English y termino en Espanol: toward a typology of code-switching. *Linguistics* 18: 581–618. https://doi.org/10.1515/ling.1980.18.7-8.581

Poplack, S., and Sankoff, D. (1984). Borrowing: the synchrony of integration. *Linguistics* 22: 99–136. https://doi.org/10.1515/ling.1984.22.1.99

Salem, F. (2017). *The Arab Social Media Report 2017* (Arab Social Media Report Series, Vol. 7). Dubai: Governance and Innovation Program, MBR School of Government.

Spiro, S. (1895). *An Arabic-English dictionary of the Colloquial Arabic of Egypt*. Beirut: Librairie du Liban.

Spiro, S. (1904). *Note on the Italian words in the modern spoken Arabic of Egypt*. Cairo: Al Mokattam Print Office.

Versteegh, K. (2009). Loan Verbs in Arabic and the DO-construction. In E. Al-Wer & R. de Jong (eds.), *Arabic Dialectology: In Honour of Clive Holes on the Occasion of His Sixtieth Birthday* (pp. 187–200). Leiden, Boston: Brill. https://doi.org/10.1163/ej.9789004172128.i-298.74

Wohlgemuth, J. (2009). *A typology of verbal borrowings*. Berlin, New York: Mouton de Gruyter. https://doi.org/10.1515/9783110219340

Woidich, M. (2006). *Das Kairenisch-Arabische, Eine Grammatik*. Wiesbaden: Harrassowitz.

Phonetical and morphological remarks on the adaptation of Italian loanwords in Libyan Arabic

Luca D'Anna

The University of Mississippi

The contact between Italian and Libyan Arabic, whose earliest traces date back to the first half of the XIX century, intensified in the decades immediately preceding the Italian occupation of Libya (1911). The number of Italian loanwords in Libyan Arabic can be estimated at about 700 lexical items, although for some of them the source might be another Romance language. The present study integrates the loanwords provided by Abdu (1988) with more lexical items collected from Yoda (2005), Pereira (2010) and the author's personal fieldwork. The data obtained are subsequently analyzed from a phonetical and morphological perspective, contributing to the knowledge of the processes of adaptation of Italian loanwords in Libyan Arabic.

Keywords: Libyan Arabic, Italian, loanwords, borrowings, Arabic dialectology, Arabic linguistics

Introduction

When Italy finally started the Libyan campaign, on October 4, 1911, "Libya had been an idée fixe in Italy for almost three decades" (Segrè 1974: p. 20). The Italian colonial aspirations on Libya can be traced back at least to 1881, when France occupied Tunisia and inflicted a heavy blow on Italy's foreign policy. The reasons adduced to justify the invasion were not different from the French or the English ones (the "white man's burden", Italy's natural right to the so-called *Fourth Shore* and to claim its share of the dying Ottoman Empire), but were strengthened by the alleged necessity to find a suitable outlet for the overabundance of unemployed workers, especially in Southern Italy. The following years saw the beginning of a "peaceful", mostly economic penetration, supported by the Italian Banco di Roma. This penetration paved the ground to the ultimate colonial occupation and

included, among other things, the publication and distribution of newspapers and periodicals in Italian. The first Libyan newspaper in Italian was *L'eco della stampa* (1892), followed by *Il giornale di Tripoli* in 1910. After the 1911 invasion, the colonial administration closed these newspapers and replaced them with two new bilingual (Arabic – Italian) ones, *Barīd Ṭarābulus* 'The Tripoli Post' and *Nuova Italia* (Camera D'Afflitto 2007: p. 61). The contact between Arabic and Italian in the Maghreb started with the beginning of the colonial period, since Italian immigrants had been present in Northern Africa (mainly Tunisia) since at least 1830 (Triulzi 1971: p. 154). The Italian colony in Tunisia left a linguistic trace in the many Italian loanwords that entered Tunisian Arabic and have been analyzed by Cifoletti (1994, 2004) and Airò (2007). A small Italian-speaking colony was also present in Tripoli before the invasion (Segrè 1974: p. 41). Despite this early presence, however, the colonial occupation of the country, started in 1911 but completed only in 1932, represents the turning point for the history of the contact between the two languages, because it radically changed the relation between the native population and the foreign colonizers. This does not mean that the Italian colonial administration, unlike its French counterpart, had any particular interest in spreading the Italian language or culture in the *Fourth Shore*. The new power (im)balance, however, made it necessary for Libyans to acquire some degree of knowledge of the Italian language to live under the new rulers. Colonial settings usually represent scenarios of unidirectional bilingualism, due to the power imbalance above mentioned. In such cases, speakers of the subaltern social group need to "…import into their own language word-forms acquired through interaction with group A in the relevant domains (Matras 2009: p. 58)." Borrowings, however, are rarely limited to specific vocabularies, but easily spread to the domains of grammatical words and even morphology (the Turkish agentive suffix *–ži* in Libyan and other Arabic dialects is a case in point). The reason for this kind of borrowings lies in the strict control over the selection of words that speakers of the subaltern group are required to maintain while communicating in the dominant language. Since the same tight control does not apply when communicating with fellow-speakers of the socially subaltern language, borrowing can happen on a larger scale than expected (Matras 2009: p. 59).

This is the period in which the greater part of the Italian loanwords entered Libyan Arabic, even though we do not have studies analyzing in much detail the social structure of colonial Libya and the relation between the native Libyans and the colonizers.

The contact between Italian and Libyan Arabic, however, did not cease with the end of the colonial period. Libya obtained full independence on December 24, 1951, but several thousands of Italian colonists and citizens remained in the former colony and continued to enjoy a high social status. In the summer of 1970, when the revolutionary regime led by a young Muammar Gaddafi announced the

expulsion of all the Italian citizens (who were given a three-month deadline to leave the country) and the confiscation of all their properties, twenty thousand Italians left the *Fourth Shore* (Segrè 1974: p. 181).

The first two decades of the revolutionary government were characterized by intransigent language policies. The Latin alphabet was banished in favor of the exclusive use of the Arabic one in the public sphere (e.g. road signs and official documents). Starting from 1984, moreover, no foreign language was taught in Libya for over a decade (Pereira 2008: p. 57). During this period, as was to be expected, Gaddafi's government tried to implement a strong policy of Arabization, which included, among other things, a systematic replacement of foreign (for the largest part Italian) words with Arabic counterparts. These efforts, however, almost never yielded the expected results, as already happened in other Arab countries in the post-colonial period.[1] At the turn of the twenty-first century, however, these strict policies were abandoned in the light of the better relations between Libya and Italy. In 2005, eventually, a Department of Italian studies was inaugurated at the University of Benghazi (former Garyounis University), thanks to the effort of the Department of Italian Studies of the University of Palermo, who donated books and provided instructors and professors (Roberto Sottile, personal communication).

More than one century of contact between the two languages resulted in the adoption of a high number of Italian loanwords in Libyan Arabic. Not all these words are known to the average Libyan. Some of them, mainly cultural loanwords introduced together with the new items or technologies they designed, were confined to specialized fields and are rapidly fading from use. Others, on the other hand, have resisted both the policies of Arabization and the strong competition of English (which is nonetheless gaining ground) and are still used by all Libyans. The terminology concerning cars and their parts is probably, in this respect, one of the best examples.

The great wealth of Italian loanwords in Libyan Arabic, despite being acknowledged in most studies concerning the different dialects spoken in the country, has never been the object of a thorough study. Abdu (1988) represents an invaluable source, since the author compiled a dictionary of all the loanwords that could be traced back to an Italian origin, even though, in some cases, the donor language might actually be another Romance language. For the present work, Abdu's list has been integrated with loanwords gleaned from the sources published after 1988, mainly Yoda (2005) and Pereira (2010), and collected during my personal fieldwork. The adaptation of the approximately 700 loanwords collected (for a small number of them, as said before, the donor language is uncertain) has then been phonetically

---

1.   For a survey concerning Arabization policies throughout the Arab world, see Bassiouney (2009: pp. 210–256).

and morphologically analyzed. Given the limited scope of this work, we focused our analysis on those areas in which the Libyan Arabic and Italian phonemic inventories diverge, highlighting the patterns that underlie the different strategies of adaptation. The study of nominal morphology focused on the deglutination of the article, gender of loanwords and the influence of Arabic nominal patterns (*ʔawzān*) on the adaptation of the borrowings. A brief paragraph analyzes the different strategies employed for the adaptation of verbs. The study does not claim to be exhaustive, but it aims at offering a contribution in a field in which much research is still needed.

1.   Phonetics

As previously said, the analysis of the phonetic and morphological adaptation of Italian loanwords in Libyan Arabic provided in the following paragraphs has been mainly based on the data collected by Abdu (1988). The phonetic adaptation of the Italian loanwords contained in his dictionary is, thus, to be considered as only valid for speakers of Western Libyan Arabic. Differences with Eastern Libyan Arabic and the varieties spoken in the so-called (although this classification appears to be questionable) transitional zone are, thus, to be expected, for two main reasons: (1) the obvious phonetic differences between Western and Eastern Libyan varieties (Pereira 2008: pp. 53–56) and (2) different phonetic realisations due to different input varieties of Italian. While our data concerning the phonetics of different Libyan varieties are sufficient to gauge their effect on the adaptation of loanwords, little is known about the dialectal background of the colonizers who settled in the different regions of Libya. In order to obtain a better picture of the Arabic spoken in the Tripoli region, moreover, we added to Abdu's database all the Italian loanwords found in Yoda's description of the Judeo-Arabic once spoken in the Libyan capital (whose last speakers currently live in Italy or Israel). This variety features a number of phonetic isoglosses that set it apart from the Muslim dialect and that, possibly, played a role in the different adaptation of loanwords. These two sets of second-hand data, finally, were integrated with a small number of loanwords not listed by Abdu and Yoda but collected during the author's fieldwork.

1.1   Consonants

A complete analysis of the way in which native speakers of Libyan Arabic adapted Italian loanwords probably goes beyond the scope of the present work. For this reason, we will limit our analysis to those areas where particularly interesting phenomena might be expected, either because Italian features phonemes that are absent

from the phonemic inventory of Libyan Arabic or because, on the contrary, a single Italian phoneme can be perceived as two distinct phonemes by native speakers of Libyan Arabic (among other reasons, because of a different vowel environment).[2]

1.1.1   *Italian phonemes absent from the phonemic inventory of Libyan Arabic*
The Italian phonemes absent from the phonemic inventory of Libyan Arabic are:

1. The bilabial voiceless occlusive /p/;
2. The labiodental voiced fricative /v/;
3. The alveolar voiceless affricate /ts/;
4. The alveolar voiced affricate /dz/;
5. The post-alveolar voiceless affricate /tʃ/;
6. The post-alveolar voiced affricate /dʒ/ represents a slightly different case. It is, in fact, realized as either an affricate or a fricative consonant in Modern Standard Arabic, while its *standard* realisation in Libyan Arabic is the fricative /ʒ/ (Pereira 2010: p. 62). This poses a double problem, since speakers might be, in theory, able to realize the Italian phoneme, but they might also neglect the difference between the Italian affricate and the native Libyan fricative;
7. The palatal lateral approximant /ʎ/ and the palatal nasal /ɲ/.

1.1.1.1   *The voiceless bilabial occlusive /p/*
The adaptation of this phoneme is by far the most predictable. It occurs 133 times in our data and is realized as its voiced counterpart [b] 131 times (98.5%), the only two exceptions being *pastikkāt* 'pills' and *spīritu* 'spirit', 'alcohol', probably due to the fact that they entered Libyan Arabic through a written medium.

1.1.1.2   *The voiced labiodental fricative /v/*
The adaptation of the Italian labiodental voiced fricative features, on the contrary, four possible realisations, among which only one ([v]) scores significantly higher than the others. The consonant occurs 41 times, while the four possible realisations are:

1. Preservation of the voiced labiodental fricative [v] (14 occurrences, 34.14%);
2. Labiovelar glide [w] (12 occurrences, 29.26%);
3. Voiceless labiodental fricative [f] (8 occurrences, 19.51%);
4. Voiced bilabial occlusive [b] (7 occurrences, 17.07%);

---

2.   On the role of perception in the learning process, see Best (1995). On the phonetic adaptation of Italian loanwords by native speakers of (Moroccan) Arabic, see Mori (2007).

Some of the possible realisations seem at least partially conditioned by the neighbouring phonetic environment. The devoiced realisation [f], for instance, occurs 8 times, seven of which share the following environment:

a.   [v] occurs in intervocalic position;
b.   [v] occurs in pre-tonic position;

e.g. *kašafīta* 'screwdriver', It. [katʃaˈvite]; *rišifūṭṭa* 'receipt', It. [ritʃeˈvuta]; *falīža* 'luggage'. It. [vaˈlidʒa], *fāzulīna* 'vaseline', It. [vazeˈlina].

The reduction of the voiced labiodental fricative to a labiovelar glide, instead, presents a less clear distribution. It occurs 11 times, and the environment that seems to particularly trigger this realisation is the intervocalic -ava- (6 occurrences).

e.g. *lawāžu* 'car washing', It. [laˈvadːʒu]; *lawandīnu* 'sink', It. [lavanˈdino]; *manawāli* 'unskilled worker', It. [manoˈvale].

The remaining 5 occurrences appear:

a.   When [w] is in initial position and is followed by a vowel (e.g.: *warakīna* 'bleach', It. [variˈkina]);
b.   When [w] is in final position and is preceded by a vowel (e.g.: *kāw* 'cable', It. [ˈkavo]);
c.   When [w] occurs between a vowel and a lateral approximant [l] (e.g.: *diāwlu* 'devil', It. [ˈdjavolo], *ṭāwla* 'table', It. [ˈtavola]);[3]

Based on our data, we can thus conclude that an interconsonantal environment inhibits the adaptation [v] → [w] and that the presence of a low vowel either before or after the consonant seems, on the contrary, to favor it.

The distribution of the [v] → [b] adaptation is, in its turn, not completely clear. This realisation occurs when the original [v] is preceded by an alveolar trill consonant (e.g. *sirbīs* 'service', It. [serˈvit:sjo]; *kūrba* 'curve', It. [ˈcurva]), but it also seems that the presence of another voiced bilabial occlusive in the word may trigger phenomena of progressive or regressive assimilation (e.g. *bugabāndi* 'vagrant', 'troublemaker', It. [vagaˈbondo]; *bābūr* 'steamship', 'kerosene stove', It. [vaˈpore]).[4]

Quite surprisingly, the preservation of the phoneme is the most common realisation, even though it is not part of the phonemic inventory of Libyan Arabic.

---

3.   In these two samples, the Italian original words feature an open syllable [vo], including a back vowel that disappears in Libyan Arabic but that probably influences the transition towards a labiovelar glide. For the rules governing the syllabic structure of Libyan (and other Maghrebi) dialects, see Marçais (1977: 26).

4.   In some cases, it is well possible that the loanword entered Libyan Arabic not via standard Italian, but rather via dialectal forms, particularly from Southern Italy (Sicily), already featuring the evolution [ˈkurva] → [ˈkurba] or [serˈvit:sjo] → [sirˈbid:zu].

It never occurs in interconsonantal position, while it is particularly common in intervocalic environments, especially when V_2 is not [a].

Given the number and the complexity of the variables involved, it is clear that other factors, unknown to us, should be taken into consideration, such as the medium through which the loanword was borrowed and the social environment in which the borrowing occurred.

1.1.1.3   *Affricates*

As said earlier, the phonetic inventory of Italian features four affricates: /ts/, /dz/, /tʃ/ and /dʒ/. A comparative analysis of the adaptation of /ts/ and /tʃ/ yielded similar results:

Realisation of [ts]

[ts] → 15 occurrences
[ts] → [s]: 10 occurrences   66.66%
[ts] → [sˤ]: 5 occurrences   33.33%

Realisation of [tʃ]

[tʃ]: 27 occurrences
[tʃ] → [ʃ]: 20 occurrences   74.07%
[tʃ] → [ʒ]: 3 occurrences   11.11%
[tʃ] → [tʃ]: 2 occurrences   7.40%
[tʃ] → [s]: 2 occurrences   7.40%

When the affricate is not geminate, it is never (in the case of [ts]) or very rarely (in the case of [tʃ]) preserved. Most of the times, it is reduced to its fricative or sibilant component, which in the case of [s] can also be emphasized. When the consonant is geminate, on the contrary, the patterns of adaptation change as follows:

Realisation of [tːs]

[tːs]: 8 occurrences
[tːs] → [ts]: 2 occurrences   25%
[tːs] → [st]: 1 occurrence   12.5%
[tːs] → [s]: 1 occurrence   12.5%
[tːs] → [sː]: 1 occurrence   12.5%
[tːs] → [ṣ]: 1 occurrence   12.5%
[tːs] → [sˤː]: 1 occurrence   12.5%
[tːs] → [z]: 1 occurrence   12.5%

Realisation of [tːʃ]

[tːʃ]: 8 occurrences
[tːʃ] → [tʃ]: 5 occurrences   62.5%
[tːʃ] → [ʃ]: 2 occurrences   25%
[tːʃ] → [dʃ]: 1 occurrence   12.5%

The gemination of the affricate makes it more easily perceivable by native speakers of Libyan Arabic, which increases the possibility of its being preserved. For [tːs], which was never preserved where non-geminate, the percentage of preservation is 25%, while it reaches 62.5% for [tːʃ], which was already preserved in the 7.40% of our samples when non-geminate.

The analysis of the data concerning the adaptation of the Italian [dʒ] and [dːʒ], on the other hands, yielded different results.

Realisation of [dʒ]		Realisation of [d:ʒ]	
[dʒ]: 18 occurrences		[d:ʒ]: 14 occurrences	
[dʒ] → [ʒ]: 18 occurrences	100%	[d:ʒ] → [ʒ]: 12 occurrences	85.71%
		[d:ʒ] → [ʒ:]: 1 occurrence	7.14%
		[d:ʒ] → [nʒ]: 1 occurrence	7.14%

The voiced postalveolar affricate [dʒ] is never preserved. This contradicts both the data concerning the adaptation of the other two affricates and the presence of the affricate /dʒ/ in the phonemic inventory of Modern Standard and Classical Arabic. It is thus clear that native speakers of Libyan Arabic, who already systematically realize etymological /dʒ/ as [ʒ], are not sensitive to the difference between the fricative and the affricate.[5]

1.1.1.4   *The palatal lateral approximant /ʎ/ and the palatal nasal /ɲ/*
The palatal articulation of these two phonemes is never preserved. In both cases, the pronunciation is dissimilated as a sequence formed by the regular alveolar lateral approximant [l] / alveolar nasal [n] and the palatal approximant [j], systematically yielding [lj] and [nj] respectively.

> e.g. *baṭṭalyōni* 'battalion', It. [bat:aˈʎ:one];
> *butīlya* 'bottle', It. [boˈt:iʎ:a];
> *famīlya* 'family', It. [faˈmiʎ:a];
> *mālya* 'jersey', It. [ˈmaʎ:a];
> *fūnya* 'sewer', It. [ˈfoɲ:a];
> *lazānya* 'lasagna', It. [laˈzaɲ:a];
> *skarōnya* 'bad luck', It. [skaˈroɲ:a];
> *kubbāniyya*[6] 'fellowship', It. [compaˈɲ:ia];

1.1.2   *Italian phonemes with possibly different Libyan outputs*
In this section, we will briefly discuss the adaptation of the Italian phonemes that can be perceived, and consequently adapted, as two distinct phonemes in Libyan Arabic. We will focus our attention, in particular, on the treatment of the Italian alveolar occlusive (/t/ and /d/) and sibilant (/s/) phonemes.

The Italian voiceless alveolar occlusive /t/ is realized, in Libyan Arabic, as either [t] or its emphatic counterpart [tˤ]. As the chart reported below clearly demonstrates, the emphatization process is favoured when the consonant is followed by a low vowel and (partially) inhibited by the presence of high vowels.

---

5.   Pereira (2010: pp. 64–65), however, notes that in Italian loanwords and in the Turkish suffix *-ği* the Arabic *jīm* is realized as a pre-palatal affricate [dʑ].

6.   Here /ɲ/ is followed by a stressed [i] in Italian.

The voiceless alveolar sibilant [s] shows a behaviour that closely resembles that of the occlusive [t]. The emphatization more frequently occurs in a low vowel environment, rarely when high back vowels are involved and never with high front vowels. If we take into consideration the small (and statistically irrelevant) sample of the loanwords beginning with *sa-* in the donor language, [s] is emphasized in seven out of nine occurrences.

	Occurrences of [tˤ]		Occurrences of [sˤ]	
ṭe	[ṭe], 0			
ṭo	[ṭo], 23	[ṣo]	[ṣo]; 2	
ṭi	[ṭi], 26	[ṣu]	[ṣu]; 7	
ṭu	[ṭu], 36			[ṣa], 23
ṭa	[ṭa], 124	[ṣa]		
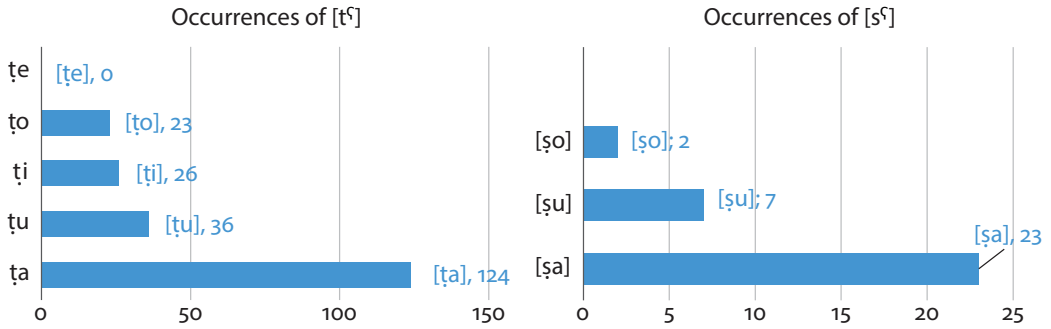
The voiced alveolar occlusive [d], on the contrary, is almost never emphasized (only three emphatic realisations occur).

1.2 Vowels

The major difference between the vowel system of Italian and Libyan Arabic is that Italian, contrarily to Libyan Arabic, does not assign any phonological value to vowel quantity. It is thus interesting to see how the two categories of stress and vowel length interact.

1.2.1 *Stressed vowels*

As largely predictable, the analysis conducted on our data reveals that stressed vowels are almost systematically perceived as long vowels.

Stressed [a]: 202 occ.		**Stressed [e]**: 124 occ.		**Stressed [i]**: 111 occ.	
[a:]: 182 occ.	90.09%	[e:]: 96 occ.	77.41%	[i:]: 104 occ.	93.69%
[a]: 14 occ.	6.93%	[i:]: 10 occ.	8.06%	[e:]: 5 occ.	4.50%
[e:]: 2 occ.	0.99%	[i]: 7 occ.	5.64%	[i]: 2 occ.	1.80%
[u:]: 1 occ.	0.49%	[a]: 4 occ.	3.22%		
[u]: 1 occ.	0.49%	[u]: 1 occ.	0.8%		
[i]: 1 occ.	0.49%				

Stressed [o]: 151 occ.   **Stressed [u]**: 41 occ.

[o:]: 93 occ.   61.58%   [u:]: 34 occ.   82.92%

[u:]: 38 occ.   25.16%   [o:]: 6 occ.   14.63%

[u]: 13 occ.   8.60%   [i:]: 1 occ.   2.43%

[o]: 3 occ.   1.98%

[a]: 2 occ.   1.32%

[a:]: 1 occ.   0.66%

[e:]: 1 occ.   0.66%

Percentages are pretty similar and show that stressed Italian vowels are realized as long vowels in Libyan Arabic, even before a triconsonantal cluster (e.g. *nāstru* 'ribbon', It. ['nastro]). In words with three or more syllables, however, a stressed vowel in the first syllable has a relatively higher probability to result in a short vowel (e.g. *maṣkara* 'mask', It. ['maskera]; *maṣtaši* 'mastic', It. ['mastitʃe];[7] *valvala* 'valve', It. ['valvola]).

1.2.2   *Unstressed vowels*

The situation concerning unstressed vowels is, as was to be expected, more nuanced, given the peculiar status of short vowels in unstressed syllables in Maghrebi dialects (Marçais 1977: p. 26). The clearest tendency emerging from our data consists in the raising of post-tonic vowels. When low and middle vowels ([a], [e], [o]) are not preserved or realized as qualitatively different middle or low vowels, in fact, the percentage of raised realisations (counting both [i] and [u]) reaches 82.78%, while lowered realisations account for the remaining 17.22%.

Table 1. Percentage of raised realisation among non-preserved low and middle vowels

	[a]	[e]	[o]
[i]	71.42	79.34	72.68
[u]	14.28	–	10.64
Tot	85.70	79.34	83.32

On the other hand, as predictable from the previous pattern, high vowels were particularly stable. The data provide only two occurrences of a lowered post-tonic [i][8] and no occurrence at all of a lowered [u].

---

7.   The realisation of *maṣkara* and *maṣtaši* is, however, also influenced by their casual resemblance to native nouns in the *mafʕal(a)* form. See 2.1.1.2.

8.   One of the two occurrences, *maṣtaši* 'mastic', It. ['mastitʃe], is doubly influenced by the presence of the preceding emphatic phoneme and by a probable adaptation to the Arabic nominal pattern $maC_1C_1aC_3a$.

It is much more difficult to recognize consistent patterns in the treatment of pre-tonic vowels. If we analyze the adaptation of words containing mid-vowels [e] and [o] in pre-tonic position, however, a tendency towards raising seems to emerge nonetheless. [o] is raised in 62.02% and lowered in 31.63%, while [e] is raised in 60.28% and lowered in 38.23%. The two possible resulting adaptations of [e] ([i] and [a]) and [o] ([u] and [a]) seem, in this case, to be in free distribution, since they occur in the same environment (e.g. *samēnsa* 'roasted seeds', It. [se'mentsa] but *žirdīna* 'garden, public park', It. [dʒar'dino]).

In some cases, the adaptation might be following rules of vocalic harmony (e.g. *ṣimyāṣ* 'axle', It. [semi'ase]; *marcānti* 'merchant', It. [mer'kante]; *madālya* 'medal', It. [me'daʎa]), but the number of counterexamples (e.g. *bansyōn* 'pension', It. [pen'sjone]) is too high to allow generalisations. In the preceding samples, for instance, the influence of Arabic nominal patterns cannot be excluded.

A more marked tendency towards vowel harmony can be detected with regard to the realisation of pre-tonic [o], even though our data feature a (smaller) number of counterexamples also in this case (e.g. [o] → [a]: *kanatēra* 'singlet', It. [kano't:jera]; *žakkatōri* 'soccer player', It. [dʒoka'tore]; *manawāli* 'unskilled worker', It. [mano'vale]; *barkamazērya* 'damn!', It. ['porka mi'zerja] but *burṭalāmba* 'lamp holder', It. ['porta 'lampada]; [o] → [i]: *kumidīna* 'night stand', It. [komo'dino], *ṭizzīna* 'dozen'; It. [do'd:zina]; *similīna* 'semolina', It. [semo'lino]).

1.3   Assimilation and dissimilation

The adaptation of Italian loanwords in Libyan Arabic often gives origin to phenomena of dissimilation. Apart from the dissimilation of affricates (see 1.1.1.3), geminate consonants are frequently dissimilated, as the following examples clearly show (e.g. *burṭaškubba* 'baseboard', It. [bat:i'skopa], *birmēstu* 'permission', It. [per'mes:o]; *bēsta* 'rag', It. ['pet:sa]; *rānžu* 'ray' ['rad:ʒo]). It is interesting to note that the dissimilation always results in a consonantal cluster.

In our data, moreover, sequences of syllables containing [l] and [r] are not allowed in the same word. In such cases, when the original word contains a sequence of two [l], the first is dissimilated in [r]. When, on the contrary, the word features two [r], it is the second one that is dissimilated in [l]. It seems, in other words, that whenever two [r] or two [l] appear in the same word, a dissimilation process is triggered that results in a word featuring [r] in the first syllable containing one of the two phonemes and [l] in the second (e.g. *skarbēllu* 'chisel', It. [skal'pel:o]; *barawolṭi*

'bumper',[9] It. [para'urti]; *burtēli* 'goalkeeper', It. [por'tjere]; *rigōli* 'penalty kick', It. [ri'gore], *varvəli* 'valves', It. ['valvole]).

Assimilation is comparatively much rarer. The only occurrence in our data is *kubbāniyya* ('company', It. [kompa'ɲia]), in which the bilabial nasal [m] is assimilated to the followed voiced bilabial occlusive [b].

2.   Morphology

This section will concisely analyze the morphological aspect of the integration of loans.

2.1   Nouns

The integration of borrowed nouns in Libyan Arabic features some interesting phenomena that will be briefly illustrated with reference to deglutination, gender and the underlying influence of Arabic nominal patterns.[10]

2.1.1   *Deglutination of the article*
Deglutination of an initial syllable wrongly perceived as a definite article is a contact-induced phenomenon widely attested since the first stages of the Arabic language in the diaspora. Names of famous urban centres, such as *al-Iskandariyya* (< Gr. Ἀλεξάνδρεια), clearly show that initial syllables of the type *vl-* were often deglutinated because of their formal resemblance with the Arabic definite article *al-* (*əl-* or *il-* in most forms of spoken Arabic) (Cifoletti 2008: p. 128). This phenomenon is particularly widespread also among the Italian loanwords, as the following samples clearly demonstrate:

1.   *bērgu* 'hotel', It. [al'bergo];
2.   *ʔanṣūla* 'sheet', It. [len'tswolo];
3.   *ʔastēk* 'elastic', It. [e'lastiko];

The [l] of the Arabic definite article, moreover, is subject to regressive assimilation if followed by one of the so-called 'solar letters' (interdental, alveolar and post-alveolar phonemes). This results in the spread of the deglutination process to words beginning with a *vC-* syllable, where *C* belongs to one of the aforementioned categories:

---

9.   For this word, however, the non-dissimilated forms *baṛawūrti* and *baṛawūṭ* (with deletion of the second [r]) also occur.

10.   For a concise analysis of plural forms, see Abdu (1988: 268–269).

4. *čāyu* 'steel', It. [aˈtːʃajo];
5. *šēta* 'hatchet', It. [aˈtːʃetːa];
6. *takku* 'attack', It. [aˈtːakːo];

While the loanwords reported above all start with traditionally 'solar' letters, it is not uncommon to find deglutinated words beginning with 'lunar' letters (i.e. consonants that do not assimilate the [l] of the article in Modern Standard or Libyan Arabic):

7. *bandašīti* 'appendicitis', It. [apːendiˈtʃite];
8. *byānti* 'plant, structure', It. [imˈpjanto];
9. *fīšu* 'office', It. [uˈfːitʃo].

These instances should be read as indicative of a general tendency towards a spread of the phenomenon to words not beginning with syllabic sequences perceivable as articles. The loanword *farīnza* 'difference' (It. [difːeˈrentsa]), for example, features the deglutination of an initial *di-*, probably perceived as the Italian preposition *di* 'of', 'from'. The phenomenon, as already noted by Cifoletti (2008), probably goes beyond the simple deglutination of articles (Cifoletti 2008: p. 128).

2.1.1.1 *Gender*
Italian and Arabic share a convergence in the marking of feminine gender (mostly *-a* in both languages) that makes the preservation of the original gender in loanwords a common event. When loanwords show a different gender than the original word, thus, interferences from the adstratal language (in this case, Libyan Arabic) might be assumed. It is, for instance, interesting that the loanwords *martēlla* 'hammer' (It. [marˈtelːo]) and *mullīna* 'mill' (It. [muˈlino]) feature a final *-a*, even though post-stress [a] is almost systematically raised. It is possible, then, that the feminine gender of the two words in Arabic (respectively CA *miṭraqa* and CA *ṭāḥūna*) interferes with the adaptation of the loanwords. The same semantic interference might be hypothesized for loanwords like *mandarīna* 'tangerine' (It. [mandaˈrino]) and *bzella* 'pea' (It. [piˈselːo]). In this case, however, the interference might be due to the singulative suffix *-a* added to names of fruits and plants to obtain the so-called *ism al-waḥda* 'noun of unity' (e.g. *tuffāḥ* 'apple' → *tuffāḥa* 'an apple).

2.1.1.2 *Influence of Arabic nominal patterns*
The influence of Arabic nominal patterns (*ʔawzān*) is noticeable in the unusual phonetic realisation of a number of nouns. The interference is not systematic, but whenever an Italian noun displays a syllabic structure formally resembling an Arabic nominal pattern, this pattern exerts a sort of "attraction", often modifying the expected (based on the outcomes of nouns sharing the same phonetic features)

phonetic adaptation of the noun. In this case, Libyan Arabic usually assigns the noun to the same category of sound or broken plural as native words actually belonging to the nominal pattern.

1	*barrāka*, pl. *barārīk*	'hut, shed'	(faʕʕāla / faʕālīl)	<	It. [baˈrakːa]
2	*bedūn*, pl. *ʔubbādīn*	'can'	(faʕālīl)	<	It. [biˈdone]
3	*buṭma*	'botton'	(fuʕla)	<	It. [boˈtːone]
4	*kālīs*, pl. *kwālīs*	'type of carriage'	(faʕālīl)	<	It. [caˈlesːe]
5	*kardūn*	'joint'	(faʕūl)	<	It. [carˈdano]
6	*gallarriyya*	'gallery'	(faʕʕāliyya)	<	It. [galːeˈria]
7	*lattariyya*	'milk shop'	(faʕʕāliyya)	<	It. [latːeˈria]
8	*mikyāṭa*	'macchiato'	(mifʕāl)	<	It. [maˈkːjato]
9	*māṭūṛ*	'engine'	(fāʕūl)	<	It. [moˈtore]
10	*ṣimyāṣ*, pl. *ṣimāyyīṣ*	'axle'	(mifʕāl / mafāʕīl)	<	It. [semiˈasːe]
11	*tantūra*	'tincture of iodine'	(faʕūla)	<	It. [tinˈtura]
12	*bābūr*	'steamship, kerosene stove'	(fāʕūl)	<	It. [vaˈpore]

2.2   Integration of verbs

The study of the different strategies employed for the integration of verbs is limited by the nature of the material under analysis. The 'light verb' strategy, in fact, involving the use of an inherited dummy verb (usually 'to do' or similar verbs) plus a noun or a frozen form of the borrowed verb (Matras 2009: p. 176–181), is not likely to be recorded in a dictionary (most nouns can become verbs through this expedient). It is worth mentioning, however, that the corpus contains two interesting cases of indirect insertion, with the Italian stem only phonetically adapted to fit into the Arabic conjugation system:

1.   *ʔitāki aḍ-ḍayy* 'switch on the light', LA *attakku* 'attack', It. [atːaˈkːo];
2.   *baḷḷa / ibaḷḷi* 'to dance', LA *baḷḷu* 'ball', It. [ˈbalːo];

In this case, it is unclear whether Libyan Arabic borrowed the verb (either the stem or an imperative) or created the verb from a borrowed noun.

A strategy typically found in Semitic languages, but possibly signaling the highest level of integration, consists in the extraction of a triconsonantal / quadriconsonantal root from the original verb, then treating it as an inherited one:

1.   **f – r – m*: *firəm* 'to sign', It. [firˈmare] 'to sign';
2.   **f – r – y – z*: *mfaryiz* 'out of order', It. [ˈfwori ˈuzo] 'out of order';[11]

---

11.   In this case, the loanword is actually an active participle, which entails the existence of an underlying verbal form, however implicit.

3. *m – n – k: *mannak* 'to be missing', It. [man'kare] 'to be missing';
4. *b – w – z: *bawaz* 'to boast', LA *bōza*, It. ['poza] 'affectation';
5. *ṣ – g – r: *ṣōgir* 'to lock', LA *sigrēssa*, It. [siku'ret:sa] 'safety';[12]
6. *z – b – l: *zabbal* 'to cause trouble', LA *zbāyli*, It. ['zbaʎo] 'mistake'.

The presence of a relatively high number of highly integrated verbs, in which a root has been extracted and made productive in the morphological system of Libyan Arabic, poses interesting questions concerning the degree of bilingualism and codeswitching in which such forms were first originated.

Conclusions

Although the greatest part of our data concern the adaptation of Italian loanwords in the region of Tripoli, the previous analysis allows us to draw some conclusive remarks. As said in the introduction, the Italian and Libyan Arabic phonemic inventory diverge in many respects, which results in interesting strategies of adaptation. Some phonemes did not pose particular problems and were adopted, although marginally, in the Libyan system. /v/, for instance, is preserved in the 34.14% of the samples collected. Others, such as /p/, /ʎ/ and /ɲ/, seemed to cause insurmountable problems to Libyan speakers and were systematically adapted to fit in the Libyan phonemic inventory (in these cases, /p/ → [b], /ʎ/ → [lj] and /ɲ/ → [nj]). The case of the affricates /ts/ and /tʃ/, on the other hand, offers some insight in the way in which second-language phonemes are perceived and processed. Non-geminate affricates, in fact, are almost never perceived and reproduced. Their adaptation systematically involves dissimilation into their fricative/sibilant component. When the affricate, on the contrary, is geminate, it is more easily perceived and reproduced by speakers, with a rate of preservation that rises from 0% to 25% for /ts/ and from 7.40% to 62.5% for /tʃ/.

As far as vowels are concerned, the Italian stress is systematically perceived in terms of vowel length. This results in a relative stability of stressed vowels, while unstressed ones show a different treatment, probably conditioned by the generally precarious status of short vowels in unstressed syllables in the majority of Maghrebi dialects. Post-tonic vowels tend to be raised. When [a], [e] and [o] are not preserved, in fact, the percentage of raised realisations reaches 80%. The treatment of pre-tonic ones is more problematic and, despite a certain tendency to vowel harmony, needs more study.

---

12. This verb also gave origin to a part participle *mṣōgar* 'certified'.

From a morphological perspective, loanwords beginning with *vl-* syllables undergo a process of deglutination of the first syllable, perceived as an equivalent of the Arabic definite article. This phenomenon is widely attested since the earliest stages of Arabic and, as also pointed out by Cifoletti (2008), goes beyond the definite article. Our data include, for instance, the loanword *farīnza* 'difference' (It. [dif:e'rentsa]), in which the deglutinated syllable was obviously not perceived as a definite article, but more probably as the Italian preposition *di* 'of', 'from'.

The gender of Italian loanwords is usually preserved, probably because of the convergence of Italian and Arabic in marking (most) feminine nouns with a final *-a*. The relatively few samples in which a shift from masculine to feminine occurs, thus, are particularly interesting. They can be explained on the basis of the parallel employment, in Arabic, of the final *-a* as a singulative morpheme suffixed to collective nouns to refer to a single item (*tuffāḥ* 'apples' → *tuffāḥa* 'one apple').

This is not the only interference of Arabic patterns on the adaptation of Italian loanwords. A number of unusual adaptations, involving gender or stress shift or the adoption of broken plurals, can be explained with the influence of an underlying Arabic nominal pattern.

As far as verbal morphology is concerned, finally, the nature of our data imposes some limitations to our analysis. Light verb strategy, involving the employment of a dummy verb (such as *dār idīr* 'to do') followed by a noun, in fact, is unlikely to be recorded in a dictionary. Our data include, however, samples of indirect insertion and, above all, of root extraction. Similar phenomena occur only when the degree of contact/bilingualism is extremely high, which raises questions on the social context in which such loanwords entered Libyan Arabic. Generally speaking, this is the field in which more in-depth studies would be greatly needed, since the details of the social structure and the patterns of coexistence between colonized and colonizers are still, for the greatest part, unknown to us.

References

Abdu, H. R. (1988). *Italian loanwords in colloquial Libyan Arabic as spoken in the Tripoli region.* Unpublished PhD dissertation, University of Arizona.

Airò, B. (2007). Vecchi e nuovi italianismi nell'arabo tunisino. In M. Moriggi (ed.), *XII Incontro Italiano di Linguistica Camito-semitica (Afroasiatica).* Soveria Mannelli: Rubbettino. 169–175.

Bassiouney, R. (2009). Arabic sociolinguistics. *Topics in Diglossia, Gender, Identity and Politics.* Washington D.C.: Georgetown University Press.

Best, C. T. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange (ed.), *Speech perception and linguistic experience: issues in cross-language research* (pp. 171–204). Baltimore: York Press.

Camera D'Afflitto, I. (2007). *Letteratura araba contemporanea. Dalla* naḥdah *a oggi*. Roma: Carocci.

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

Cifoletti, G. (2004). Il plurilinguismo a Tunisi. In Bombi, R. and Fusco, F., *Città plurilingui, Lingue e culture a confronto in situazioni urbane*. Udine: Forum.127–132.

Cifoletti, G. (2008). Sur la diffusion des emprunts européens en arabe modern. In S. Procházka and V. Ritt Benmimoun (eds.). *Between the Atlantic and Indian Oceans, Studies on Contemporary Arabic Dialects*. Münster – Wien: LIT – Verlag. 125–128.

Marçais, P. (1977). *Esquisse Grammaticale de l'Arabe Maghrébin*. Maisonneuve: Paris.

Marçais, P. (2001). *Parlers arabes du Fezzân, Textes, traductions et éléments de morphologie rassemblés et présentés par Dominique CAUBET, Aubert MARTIN et Laurence DENOOZ*. Genève: Librairie DROZ.

Matras, Y. (2009). *Language Contact*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511809873

Mori, L. (2007). *Fonetica dell'italiano L2. Un'indagine sperimentale sulla variazione nell'interlingua dei marocchini*. Roma: Carocci.

Pereira, C. (2008). Libya. In K. Versteegh (ed.), *Enciclopedia of Arabic Language and Linguistics* vol. III. Leiden – Boston: Brill. 52–58.

Pereira, C. (2010). *Le Parler arabe de Tripoli (Libye)*. Zaragoza: Istituto de Estudios Islámicos y del Oriente Próximo.

Segrè, C. G. (1974). *Fourth Shore. The Italian Colonization of Libya*. Chicago – London: The University of Chicago Press.

Triulzi, A. (1971). Italian-speaking communities in early nineteenth century Tunis, in *Revue de l'Occident musulman et de la Méditerranée*, n. 9, 153–184.

Yoda, S. (2005). *The Arabic Dialect of the Jews of Tripoli (Libya), Grammar, Text and Glossary*, Verlag-Wiesbaden: Harrassowitz.

An assessment of the Arabic lexical contribution to contemporary spoken Koalib

Nicolas Quint

LLACAN-UMR8135 (CNRS/INALCO/Université Sorbonne Paris Cité)

The present paper deals with the lexical contribution of Arabic, the dominant language of Sudan, to Koalib, a Kordofanian language traditionally spoken in the northeastern part of the Nuba Mountains (Southern Kordofan, central Sudan). The study is based on a corpus of 400 Koalib items borrowed from Arabic, the main characteristics of which (social context, phonology, part of speech and semantics) are successively examined and discussed. The conclusion summarizes the main typological implications of the Arabic influence upon the Koalib grammatical system.

Keywords: borrowings, Koalib, Kordofanian, language contact, lexicon, Nuba Mountains, Sudanese Arabic

Introduction

Koalib is a Kordofanian language (Heiban branch) spoken by approximately 100,000 people living in or originating from an area including the cities of Abri, Delami, Dere, Tongole, Umm Berembeita, Umm Heitan and their surroundings, in the Sudanese province of South Kordofan (also known as the Nuba Mountains). Koalib has been in contact with Arabic for at least 250 years, since the arrival of Arabic-speaking nomads in the North-eastern parts of the Nuba Mountains and the development of the Muslim kingdom of Tegali, in the immediate vicinity of the Koalib area (Ewald 1990; Stevenson 1984).

This prolonged contact has led to the integration of an important number of Arabic-derived lexical items into contemporary Koalib, even in the most conservative (ie. least exposed to Arabic) varieties of the language. In this paper, I focus on this Arabic lexical contribution to contemporary spoken Koalib (namely the central variant of the language, called Rere in English and *ŋèréɛɽè* in Koalib), in particular on those items that are Arabic-derived and have been fully integrated

into Koalib. In section one, I introduce the reader to basic notions about the nature of the contact between Koalib and Arabic in the course of history and the lexical corpus of cca. 400 Koalib items borrowed from Arabic on which I base my observations. Section two is devoted to the way Arabic items are actually integrated into Koalib on a phonological level, as Koalib and Arabic display striking differences as regards their phonic systems. Section three will deal with parts of speech: which parts of speech are more easily borrowed, how and why, and how one and the same Arabic root can give rise to several related Koalib items belonging to different parts of speech. Section four will be concerned with semantics, trying to ascertain which semantic fields are favored in the borrowing process. Finally, I conclude, trying to relate the observations I have made about Arabic loanwords in Koalib with the general studies available on language contact and lexical borrowing (e.g. Thomason and Kaufman 1991; Winford 2003) and reflecting on the role of Arabic in the development and evolution of the lexicon of an indigenous language of the Nuba Mountains, such as Koalib.

1.   Koalib and Arabic in contact: Some basic notions

As was said in the introduction, Koalib (Kb.) and Arabic (Ar.) have been in direct contact for at least two centuries and a half, since the times of the Kingdom of Tegali, through whose influence the Koalibs first began to get in touch with Arabic and Muslim culture.

1.1   Linguistic characteristics of the contact

1.1.1   *Varieties of Arabic involved in the contact*
At least three different varieties of Arabic are of have been in contact with Koalib (Quint 2014: pp. 124–126):

i.   Standard (= Khartoum) Sudanese Arabic (Sd. Ar., based on Khartoum variety), spoken by a majority of Sudanese, is nowadays by far the Arabic variety which exerts the strongest influence upon Koalib, to which it has provided *inter alia* many terms linked with modernity (see also **4.1** below):

(1) Kb. *tèáarà* '*plane*' < Sd. Ar. *ṭayyaara(t)*.

ii.   nomadic varieties, spoken by the various groups of herders (such as the Baggara (Bagg.) and Shanabla) who drive their flocks across the Koalib area in accordance with the rhythm of their seasonal moves. Historically, the nomadic varieties were the first to be in contact with Koalib and some Koalib items can clearly be traced to these:

(2)   Kb. *àlmántàr* 'mirror' < Bagg. Ar. *al-mánḍar* ≠ Sd. Ar. *miráaya(t)*.

(3)   Kb. *Áacè*[1] [áaʒè] 'feminine proper noun' < Bagg. Ar. *ʕáaʃe*[2] ≠ Sd. Ar. *ʕáaʔiʃa(t)*.

iii.   an Arabic-based local vehicular also seems to have played a role in the dissem-ination of Arabic culture among Koalibs. Although this vehicular is no longer spoken, it is still possible to find some of its fossilized remnants in various Koalib folk-stories (Quint 2014: p. 125; 2010a: pp. 55–56).

1.1.2   *Varieties of Koalib involved in the contact*

There are several local varieties of Koalib (Quint 2009: pp. 20–23; 2006: pp. 23–25), which have been influenced in different ways by Arabic, e.g. the Northern and Western parts of the Koalib country (respectively home to the *lŭkrèŋ* and *lèmrέ* sub-tribes of Koalibs, each of which speaks its own Koalib variant) have been generally more exposed to Arabic culture than other Koalib regions. However, the available data do not allow us to compare the influence exerted by Arabic on each local Koalib variety. Rather, I will focus herein on one particular Koalib variety, namely Rere (Kb. *ŋèrέɛɽè*). Rere is traditionally spoken in and around Abri (Kb. *Âbɽè*), in the very center of the Koalib-speaking area. It is the Koalib variety I know best and which I have documented in more detail. For historical reasons,[3] Rere is also the basis of standard written Koalib: at least two versions of the New Testament have been published in Rere (NT1, NT2), as well as parts of the Old Testament (A3), some textbooks for foreigners (A1) or for children (A2), several folkstories (e.g. A5, KO) and calendars (e.g. A4, CA1), etc. In other words, there is a written corpus of Rere texts of over 1,000 pages, which can be used as a complementary source to spoken Rere in order to check the lexical influence exerted by Arabic upon the language. Accordingly to what has just been said, the label 'Koalib' will be used henceforth with the meaning of 'Rere Koalib'.

---

1.   Koalib items are normally provided under their orthographic form (for more details, see Quint 2009: pp. 169–187; 2006: pp. 189–210). When the employed orthography differs markedly from IPA conventions and when required by the discussion, a phonetic transcription is added.

2.   The pronunciation -[e] of etymological pan-Arabic and Semitic final -/a(t)/ (feminine suffix marker) is regular in Kordofan Arabic when a high vowel – here the /i/ of classical and Sudanese Arabic *ʕáaʔiʃa(t)* – precedes the suffix -/a(t)/ in the etymon (Manfredi 2010: pp. 230–231).

3.   The first Christian Missionaries who came in the 1920s to the Koalib area settled in Abri and set up to translate portions of the Bible in this variety (Quint 2009: pp. 12–13; 2006: p. 15, and references therein).

1.2   Social characteristics of the contact between Koalib and Arabic

Today, the contact between Koalib and Arabic is clearly unbalanced from a socio-linguistic point of view:

–   Arabic is the main official language used in the Sudan as a whole and in the government-held areas[4] of the Nuba Mountains in particular.[5] It is used as the sole medium in public schools as well as for any administrative matter. Arabic is also the only vehicular language used throughout the Nuba Mountains, including in SPLA-N-held areas. Khartoum Sudanese Arabic (the most powerful variety) is spoken natively or as a second language by at least 30 million Northern Sudanese. Save for some few Arabic-speaking nomads who grew among Koalib communities before the second Sudanese civil war, virtually no person having Arabic as their mother tongue and not being an ethnic Koalib is able to speak Koalib with some degree of fluency.

–   Koalib has no official status. In government-held areas, it is not supported by the administration. In SPLA-N-held areas, its spoken use is encouraged but it is not taught at school. Koalib is generally not a vehicular language (although some neighboring tribes such as Laros and Shwais sometimes use it at church services for lack of available Bible translations in their own tongues) and is rarely learnt by non Koalibs. As mentioned above, Koalib has no more than 100,000 speakers (i.e. several hundred times fewer than Sudanese Arabic). Today, ca. 50% of ethnic Koalibs speak Arabic as a first language: this applies to Koalibs who were born in big Sudanese cities outside the Koalib area but also to Koalibs who were born and raised within the Koalib area itself (where Koalib nevertheless remains the first language of many children). Except for some young children and elders, almost all Koalib-speakers are bilingual in Sudanese Arabic and code-switching between Koalib and Arabic is quite developed even during in-group exchanges. In order to discuss some topics (such as politics), most Koalib-speakers switch spontaneously to Arabic only to come back to Koalib when dealing with more traditional activities (such as cooking or farming).

---

4.   Since the second Sudanese civil-war (1983–2005), the Koalib area is cut into two: (i) one part is controlled by the Sudanese government, while (ii) the second part is under SPLA-N (Sudan People's Liberation Army-North) control. The SPLA-N is the Northern Sudanese branch of the SPLA, the ruling party of Southern Sudan, which fought against the Sudanese government during the second Sudanese civil-war. In the Nuba Mountains (including the Koalib area), a significant proportion of the population sided by the SPLA during the war and, when fighting resumed in this region after South Sudan achieved its independence in 2011, the SPLA-N pursued the fight of the SPLA in the Nuba Mountains.

5.   In the SPLA-N-held areas, English is the main administrative and school language.

1.3   The Arabic loanwords discussed in this study

If one takes into account Arabic-Koalib code-switching, virtually any Arabic word can be used in Koalib. The present paper has a narrower scope: it focuses on Arabic-derived items (i) attested in Koalib conservative speech such as folk-stories or conversational settings dealing with traditional topics which are generally discussed mostly in Koalib by Koalib-speakers and (ii) which show at least some degree of phonetic integration into the Koalib sound system. The Arabic loanwords answering (i) and (ii) represent a body of ca. 400 items, i.e. 7% of the 5,900 entries of the Koalib dictionary I have presently compiled.

2.   Phonological integration of Arabic borrowings into Koalib

2.1   Segmental integration

As the sound system and phonemic inventory of Koalib and Arabic differ in many respects,[6] Arabic borrowings generally have suffered considerable changes when being integrated into Koalib:

i.   deletion. Some Arabic consonants which are absent from the Koalib sound system have simply been dropped:

   (4) Sd. Ar. /ʕ, ʔ, h, ḥ/ > Kb. ø
         Sd. Ar. *ʕarabíyya(t)* 'car' > Kb. *àràbêa*.
         Sd. Ar. *ʔusbúuʕ* 'week' > Kb. *òcbô* [òʒbô].
         Sd. Ar. *hadiyya(t)* 'present' > Kb. *ìtîɐ* [ìðîɐ].
         Sd. Ar. *ḥaláawa(t)* 'sweet' > Kb. *àláoà*.

ii.   substitution. Other Arabic consonants, also lacking in Koalib, have been substituted by a Koalib sound displaying similar acoustic features:

   (5) Sd. Ar. /x, ġ, g/ > Kb. /k/ (back articulation).
         Sd. Ar. *xafíir* 'porter' > Kb. *kàapêr* [kàavêr].
         Sd. Ar. *ġálaṭ* 'wrong (n.)' > Kb. *káràt*.
         Sd. Ar. *gálam* 'pencil' > Kb. *kálàm*.

---

6.   For a better understanding of these differences, two phonological charts, presenting respectively the vowels and consonants of Koalib, are given in Appendix 1. Regarding Arabic, the phonemes discussed in this paper are easily recognizable for anyone familiarized with Arabic phonology; non-Arabists will have no trouble in finding their description on the Web or in any specialized publication dealing with Arabic.

(6) Sd. Ar. /s, ṣ, z/ > Kb. /ʃ/ (initial/final) ~ /ʒ/ (medial)
     Sd. Ar. *suuq* 'market' > Kb. *cûk* [ʃûk].
     Sd. Ar. *ṣúlta(t)* 'strength' > Kb. *cólttà* [ʃóltà].
     Sd. Ar. *(ána) záatu* 'myself' > Kb. *(nyí) cáttòk* [ʒáttòk].
     Sd. Ar. *ġáṣib* 'by force' > Kb. *kácèp* [káʒèp].
     Sd. Ar. *ʔiblíis* 'evil' > Kb. *ɛ̀blîc* [ɛ̀blîʃ].
     Sd. Ar. *raṣáaṣ (at)* 'bullet' > Kb. *àrcâc* [àrʒâʃ].

iii. vowel harmony. Contrary to Arabic, Koalib follows a strict vowel harmony
     (Quint 2009: pp. 34–40; 2006: pp. 34–42) whereby all vowels of a given word
     can belong to only one of the two harmonic sets, i.e. /e, ɛ, a, ɔ, o/ (low set) or
     /i, ɐ, u/ (high set). These harmonic rules are regularly applied to integrated
     Arabic polysyllabic loanwords:

(7) Sd. Ar. /a…i/ > Kb. /a…e/ or /ɐ…i/. As /a/ and /i/ belong to two different har-
     monic sets in Koalib (respectively the low and the high set), when an Arabic
     word containing both these vowels is integrated into Koalib:
     (i)   either the /i/ changes to /e/ and both vowels of the resulting word belong
           to one and the same set, namely the low one: Sd. Ar. *xafíir* 'porter' > Kb.
           kàapêr [kàavêr].
     (ii) or the /a/ changes to /ɐ/ and both vowels of the resulting word belong to
           the high set: Sd. Ar. *gamíiṣ* 'shirt' > Kb. *kɐ̀mîc* [kɐ̀mîʃ].

2.2   Tonal integration

While Sudanese Arabic is a stress language,[7] Koalib is a tone language (Quint
2009: pp. 129–148; 2006: pp. 145–165), with:

i.   two basic tones, low (L, transcribed /V̀/), e.g. *ŋwɔ̀ny* 'eggs' and high (H, tran-
     scribed /V́/), e.g. *ŋwɔ́ny* 'saliva'.
ii. several contour tones, the most frequent being falling (F, transcribed /V̂/), e.g.
     ŋâo 'water' and rising (R, transcribed /V̌/), e.g. *ŋɔ̌rppà-ná* 'the following day'.

As regards stress, Arabic loanwords regularly abide by the rules given in Table 1.

---

7. Throughout this paper, stressed syllables in Arabic polysyllabic words are indicated by an
acute accent.

Table 1. Tonal rules of integration of Arabic loanwords into Koalib

Sd. Ar. etymon		Koalib loanword		
Syllable	Position	Date	Tone	Example
Unstressed (U)	Any	Any	Low (L)	(8)
Stressed (S)	Non-final	Any	High (H)	(9)
	Final	Recent	Falling (F)	(10)
		Old	High (H)	(11)

Examples:

(8)   Sd. Ar. Kb. *fallúuka(t)* [USU] 'boat' > Kb. *pèllúkkè* [LHL].

(9)   Sd. Ar. *shánṭa* [SU] 'suitcase' > Kb. *cántà* [HL].

(10)   Sd. Ar. *al-báab* [US] '(the) door' > Kb. *lèbâb* [LF].

(11)   Sd. Ar. *kitáab* [US] 'book' > Kb. *kèttám* [LH].

That *kèttám* 'book' is probably an older loanword than *lèbâb* 'door' is supported by the fact that Sd. Ar. -/b/ has been changed into -/m/ in *kèttám* but has remained -/b/ in *lèbâb*. In the same phonetic context and for the same source language, the form which is closest to its etymon (here *lèbâb*) has indeed all chances to have been integrated more recently.

3.   Arabic borrowings according to their parts of speech and their morphological characteristics

As shown in Table 2, in the above defined (see **1.3**) corpus of Arabic loanwords studied in this paper, most items (77 + 10 = 87%) are either common or proper nouns. This result is in conformity with general available hierarchies of borrowing (e.g. Muysken 1981b quoted by Winford 2003: p. 51). In this section, I will examine in turn the main morphological characteristics of the Arabic borrowings according to the part of speech they belong to in the receiver language, namely Koalib.

Table 2. Arabic loanwords into Koalib according to part of speech

Part of speech		Number of items	%
Nouns	Common nouns	310	77
	Proper nouns	41	10
Adverbs		30	7
Verbs		11	3
Others		11	3
Total		**403**	**100**

3.1   Common nouns

3.1.1   *Formal integration*

As Arabic and Koalib are typologically quite different from each other, Arabic nouns and phrases have been subject to many kinds of reinterpretations when acquiring the status of Koalib nominal items. The most common case of reinterpretation is the agglutination of the Arabic definite article *al* 'the' to the noun root it precedes to produce a unique Koalib noun:

(12)   a.   Sd. Ar. **al**-*báab* '**the** door' > Kb. *lèbâb* 'door'.
       b.   Sd. Ar. **al**-*ʔibríig* '**the** jug' > Kb. *lìbrîk* 'jug'.

Other more anecdotical examples include:

(13)   Sd. Ar. *ʕéeʃ ar-ríif* 'bread (of) the countryside', i.e. 'maize' > Kb. *cìrím* [ʃìrím].

(14)   Sd. Ar. *ʔábu ʕáʃara* 'father (of) ten', i.e. 'type of rifle which can be filled with ten bullets' > Kb. *bácrà* [báʒrà].

3.1.2   *Paradigmatic integration*

All nominal Arabic loanwords are fully integrated in the various Koalib nominal paradigms, such as:

i.   noun classes, in which the first consonant of the nominal loanword is reinterpreted as a class-marker[8] participating to two main morphological processes:

(15) Number marking:
       Sd. Ar. *gamíiṣ* 'shirt' > Kb. *kɛ̀mîc* (SG) > *yɛ̀mîc* (PL)
       (*k* (SG)/*y* (PL) class pairing, see non-borrowed Kb. *kéɐnì* 'ear' (SG) > *yéɐnì* (PL))
       Sd. Ar. *al-ʕílba(t)* 'box' > Kb. *lélbè* (SG) > *ŋwélbè* (PL)
       (*l* (SG)/ *ŋw* (PL) class pairing, see non-borrowed Kb. *lékké* 'crippled person' (SG) > *ŋwékké* (PL))
       Sd. Ar. *ṭaagíyya(t)* 'hat' > Kb. *tàkêa* (SG) > *ràkêa* (PL)
       (*t* (SG)/ *r* (PL) class pairing, see non-borrowed Kb. *táakà* 'marriage' (SG) > *ráakà* (PL))

(16) Noun derivation:
       Sd. Ar. *táajir* 'merchant' > Kb. *táajèr* 'id.' > *ŋáajèr* 'business' (abstract noun in *ŋ*-).
       Sd. Ar. *gazáaza* 'bottle' > Kb. *kècácà* 'id.' > *ṭècácà* 'small bottle' (diminutive derivation in *ṭ*-).

---

8.   For more details, see Quint (2014).

ii.   declension. Most Koalib nouns have two different forms according to their ar-
gumental role (Subject (S) or Object (O)) in the sentence and so do the Arabic
loanwords in Koalib:

(17) Sd. Ar. *al-ʔeféndi* (from Turkish) 'Mister, Master' > Kb. *lèpéntì* [lèvéndì]
'teacher (S)' > *lèpéntìɛ́* (O).
(suffixed case marking, see non-borrowed Kb. *kwîmr̥ù* 'loving person' (S) >
kwîmr̥ùɛ́ (O)).

(18) Sd. Ar. *ramaḍáan* 'fast (n.)' > Kb. *àrmèntán* (S)' > *àrmèntánɛ́* (O).
(suffixed case marking, see non-borrowed Kb. *lèbàrttáŋ* 'rain drop' (S) > *lèbàr-
ttáŋɛ́* (O))

(19) Sd. Ar. *fallúuka* 'boat' > Kb. *pèllúkkè* (S) > *pèllùkkè* (O).
(tonal case marking, see non-borrowed Kb. *ʈùkúrnè* [ʈùgúrnè] 'grandfather'
(S) > *ʈùkùrnè* (O))

3.1.3   *Introducing new distinctions into Koalib: The case of sex-based gender*
Koalib has a multiple-gender system (expressed through class prefixes) in which
the members of one and the same noun-class share in common a salient semantic
element (e.g. LONG OBJECTS (one dimension), PLANE OBJECTS (two dimensions),
LIQUIDS, HUMANS…: see Quint 2014) but there is no morphological classes oppos-
ing 'masculine/male' vs. 'feminine/female'. Arabic, in turn, has a two-gender system
based on the opposition 'masculine' vs. 'feminine' and one of the most common
feminine markers is the suffix *-a(t)*. In one case at least, Koalib has integrated a full
masculine/feminine pair from Arabic:

(20)   Sd. Ar. *xaddáam* '(male) servant' vs. *xaddáama(t)* '(female) servant' > Kb. *kad-
dâm* (M) vs. *kàddámà* (F). Both nouns have the same object form (*kàddàmà*).

In this case, some Arabic noun gender morphology has managed to make its way
into Koalib but, at least in conservative speech, this example remains an exception.

3.2   Proper nouns

Borrowed Arabic proper nouns regularly fit into the Koalib morphological pattern
for this category; in particular they also combine with the specific object marker
ŋwÓ[9] [ŋwú (HS), ŋwó (LS)] used with any proper noun:

---

9.   Due to the above described (**2.1**, (iii)) rules of Vowel Harmony, each Koalib syllabic affix
has two forms, one for the high set (HS) and the other one for the low set (LS). The vowel of the
general form of the suffix is given with the form of the low set in uppercase (e.g. *-ŋwÓ* above).

(21)   Sd. Ar. *Már(i)yam* 'Mary' > *Mérìèm* '*Mary*' (S) > *Mérìèmŋwú* (O).

(22)   Sd. Ar. *Máṣir* 'Egypt' > *Mócòr* (S) > *Mócòrŋwó* (O).

3.3   Verbs

As Koalib verb morphology (mainly based on movement with many tonal distinctions, see Quint 2010b: pp. 296–297 (ft. 2)) is very distinct from Sudanese Arabic morphology (based on aspect and segmental affixes), relatively few Arabic verbal items have been integrated into Koalib:

(23)   Sd. Ar. *gála* (PFV)/ *yágli* (IPFV) 'fry' > Kb. *èklí* [ὲglí] 'roast (coffee)'.

(24)   Sd. Ar. *gárrab* (PFV)/ *yigárrib* (IPFV) 'draw near' > Kb. *kòrbé* 'tie (v.)'.

Some few Arabic-derived verbs have even combined with Koalib verb extensions:

(25)   Sd. Ar. *támma* (PFV)/ *yitámmi* (IPFV) 'complete' has not been integrated as such in Koalib but has given two different Koalib verb extensions:
> Kb. *tèmmàcé* 'complete' (transitive *-ÀcÉ* verb extension).
> Kb. *tèmmànné* 'be completed' (passive *-ÀnnÉ* verb extension).

(26)   Sd. Ar. *ʔámal* (PFV)/ *yaʔmáli* (IPFV) 'do'
> Kb. *áml(èt)à* [ámlèðà] 'prepare sth. (food)' (malefactive[10] *–ÀtÀ/ÈtÀ* verb extension).
> Kb. *ámlàccé* [ámlàccé] 'prepare sth. to s.o.' (benefactive *–ÀccÉ* verb extension).

3.4   Adverbs

3.4.1   *Classical loanwords*
These Koalib adverbs are generally derived from very frequent Arabic items:

(27)   Sd. Ar. *yímkin* (verbal form) 'it may be, perhaps' > Kb. *émkkìn*.

(28)   Sd. Ar. *ġásib* > KB. *kácèp* 'by force, against one's will'.

3.4.2   *Arabic-Koalib adverbial bases used in conjugation*
This is probably the most spectacular consequence of the contact between Arabic and Koalib. Indeed, Koalib has developed an original series of adverbs recycling Arabic consonantal roots and associating them with (i) recurring sets of vowels (most frequently A-E-A for trisyllabic Arabic-Koalib adverbs) and (ii) a regular

---

10.   'Malefactive' is the general label for this extension (Quint 2010b). In this specific case, *áml(èt)à* has not any MALEFACTIVE value. It is merely an applicative.

tone pattern (HB(B)). The typical form of these blended adverbs is (C)ÁC(C)ÈCÀ. At least 16 such items have been recorded (i.e. half of the Koalib adverbs borrowed from Arabic, see full list in Appendix 2), which seems to suggest that this morphological construction is productive. In most cases, the Arabic-Koalib adverb (AKA) does not seem to be derived from any specific Arabic item associated with the Arabic consonantic root.[11] Rather, it really is produced by the Koalib system itself using the Arabic consonants and the semantics associated with one specific lexical root and making them fit into a specific morphological template (not attested for Koalib lexical bases).

These adverbs are typically used to produce compound verbs, in which (i) the AKA provides the meaning and (ii) a light verb (*àaré* 'say, do' or its verbal extensions: *èerìnní* (passive) 'be said', *éccè* (benefactive) 'say to') provides the verbal inflection:

(29) Sd. Ar. *sfr* [triconsonantal root] 'travelling', e.g. *safar* (n.) '*travel*' > Kb. *(àaré) cápèrà* [ʃávèrà] 'say travelling', i.e. 'travel'.

(30) Sd. Ar. *krmʃ* [quadriconsonantal root] 'wrinkling', e.g. *itkármaʃ* (PFV)/ *yitkármiʃ* (IPFV) 'get wrinkled (cloth)') > Kb. *(àaré) kérmìcè* [kérmìʒè] = 'say wrinkling', i.e. 'be/get wrinkled (cloth)'.

In several cases, one of the consonants of the Arabic root is not maintained in the AKA, as it does not belong to the Koalib sound system (see **2.1**):

(31) Sd. Ar. *ʕlm* [triconsonantal root] 'teaching', e.g. *ʕállama* (PFV)/ *yiʕállim* (IPFV) '*get wrinkled (cloth)*' > Kb. *(àaré) állèmà* 'say teaching', i.e. 'teach (sth.)'. In this case, the initial ʕ of Arabic is dropped in Koalib.

(32) Sd. Ar. *nɟħ* [triconsonantal root] 'succeeding', e.g. *nájaħ* (PFV)/ *yínjaħ* (IPFV) '*succeed*' > Kb. *(àaré) ánycà* [áɲɟà] '(say) succeeding', i.e. 'succeed'. In this case, the final *ħ* of Arabic is dropped in Koalib.

Interestingly enough, in some cases, an AKA can coexist in Koalib with other Arabic loanwords derived from the same root:[12]

(33) Sd. Ar. *ħkm* 'judging'
> *ákkòmà* (AKA), 'judging, in a judiciary way'
+ *àkkèmé* (v.) 'judge' < Sd. Ar. *ħákam*
+ *ókkòm* (n.) 'judgement' < Sd. Ar. *ħúkum*

---

11. However, this is not an absolute rule: in (31), the Koalib AKA *állèmà* is clearly derived from the Arabic verb *ʕállama*.

12. See Appendix 2 for more examples.

3.5   Other parts of speech

They are much rarer, and comprise:

i.   conjunctions:

(34) Sd. Ar. *ya* 'or' > Kb. *yâ*.

ii.   interjections:

(35) Sd. Ar. *dúnya* 'earthly existence, world' > Kb. *dúnɪ̀ɐ* 'alas, my God'.

(36) Sd. Ar. *wa lláahi* 'by God' > Kb. *àllâe* 'alas'.

iii. prepositions:

(37) Sd. Ar. *ʔílla* 'except' > Kb. *íllì ~íllɪ̀*:

Lìcì	*lèm-èelà*	*Kálkè*	*ṭàṭṭàp*	**íllì**	*Kwɔ́ccɔ̀*	*dâk.*
people	PRF-come	Delami	all	except	Kwɔ́ccɔ̀	only

'Everybody has come to Delami[13] except Kwɔ́ccɔ̀.'

Although few Arabic prepositions have entered Koalib lexicon, it must be noted that, should their number increase, such elements would lead to typological changes in Koalib for most Koalib endogenous adpositions are postpositions: *ɛ̀ɛtúmɛ̀-kùttú* [termitary.O-under] '**under** the termitary', *ŋáo-ná* [water-in] '**in** the water'.

4.   Some semantic characteristics of Arabic borrowings

4.1   Typical semantic fields

Like in other situation of unbalanced bilingualism, Koalib is particularly prone to borrow Arabic words belonging typically to some specific semantic fields:

–   all Koalib names of the days of the week are Arabic-derived, probably because there was no such habit of naming days in traditional Koalib culture:[14]

(38) Sd. Ar. *sábit* 'Saturday' > Kb. *cébìt* [ʃébìt].

–   Muslim names (linked with Arabic culture):

(39) Sd. Ar. Muḥámmad > Kb. Mɛ́mmɛ̀t.

---

13.   Delami is one of the main Koalib cities. It is situated North of the Koalib area. *Kwɔ́ccɔ̀* is the traditional Koalib name of the 5th boy in a family.

14.   Calendar terms (and other time reference words) are commonly borrowed from dominant languages by minority languages: see Nunez (2015: p. 83–106).

(40) Sd. Ar. Al-ḥajj > Kb. Àalâny.

(41) Sd. Ar. Zéenab > Kb. Cénàp [ʃénàp].

– Christian[15] vocabulary (attested in Bible translations (NT1, NT2, A3)):

(42) Sd. Ar. kaníisa 'church' > Kb. kènícè [kèníʒè].

(43) Sd. Ar. maláak 'angel' > Kb. mèlékkà.

– items linked with modernity and new technologies, which often reach the Koalib-speaking community through Arabic:[16]

(44) Sd. Ar. *wáraga* '(sheet of) paper' > Kb. *wérkè* [wérgè].

(45) Sd. Ar. *gazáaza* 'bottle' > Kb. *kècácà* [kèʒáʒà].

(46) Sd. Ar. *kúbri* 'bridge' > Kb. *kwúbrì*.

– counting system: there is a Koalib traditional vicesimal counting system which allows its user to count until 400 (20x20). However, for bigger numbers, there does not seem to be attested forms in Koalib. Therefore, such numbers are systematically borrowed from Arabic:

(47) Sd. Ar. *álif* '1,000' > Kb. *álèp*.

As a matter of fact, in today's spoken Koalib, even in conservative speech, only the smallest numbers (under 10 or 20) are said in Koalib. All other numbers are directly borrowed from Arabic without any phonetic integration.

4.2   Conventionalized calques

Beyond lexical borrowings, Koalib (in particular the written form of the language) displays structural calques in which a lexically Koalib phrase follows in reality the syntax of Arabic instead of spoken Koalib:

(48) *lóomór lèʈʈè lètè ŋwòomòr*
period   one   of   periods
'one fine day, once upon a time' (A5: 1)

All four words of this Koalib phrase are non-borrowed. However, they correspond almost literally to the semantically equivalent written Arabic phrase:

---

15. A majority of Koalibs (roughly 60%) are Christian. There is also a strong Muslim community and some adepts of animism.

16. Historically, other languages (in particular Nubian) have played the role of passing new cultural items onto Koalib. Since the nineteenth century, English also contributes to the lexical expansion of Koalib (see also Quint 2014: p. 126–127).

(49) *fii yáwmin min   al-ʔayyáami*
      in day   of/from the-days
      'one fine day'

The proof that (48) is a calque from (standard) Arabic is that, in spoken Koalib, 'one fine day' is normally said *láamén lèʈʈè* 'day one', i.e. 'one day'. The use of (48) is therefore typical of written Koalib, which is sometimes dependent on the standard provided by Arabic, which is the first and main language most Koalibs study in Sudanese schools (see above **1.2**).

Such calques from Sudanese or standard Arabic are probably quite numerous in both written and spoken Koalib. They provide an interesting field for further studies on the influence Arabic exerts on Koalib.

5.   Conclusion

Throughout this paper, we have examined the Arabic lexical contribution to the most conservative forms of Koalib. In these Koalib lects, the Arabic element consists mostly of individual words[17] while structural borrowings and calques seem to be relatively limited (or perhaps not identified yet …). However, in the course of over two centuries of linguistic contact, this Arabic element has had amply enough time to interact with Koalib and contribute to develop original structures, the most spectacular of which being probably the Arabic Koalib adverbs studied in **3.4.2**. Other points too (such as the appearance of a male/female lexical pair (**3.1.3**) or the increasing use of prepositions (**3.5**)) are worthy of attention and may be considered as a first step towards a process of metatypy, whereby Koalib would little by little change its typological characteristics under the influence of Arabic. At any rate, the impact of Arabic is much more important in other Koalib varieties, and in particular in spoken contemporary Koalib: it is hoped that further studies will help assess the exact importance of this impact across the diverse lects and usages of today's Koalib.

Acknowledgments

---

17. See Thomason & Kaufman (1991: p. 37): "Invariably, in a borrowing situation the first foreign elements to enter the borrowing language are words".

Abbreviations

ᴀᴋᴀ	Arabic-Koalib Adverb	ᴏ	Object
Ar.	Arabic	ᴘꜰᴠ	Perfective
Bagg.	Baggara	ᴘʀꜰ	Perfect
ꜰ	Falling	ᴘʟ	Plural
ʜ	High	ʀ	Rising
ʜꜱ	High Set (Vowel Harmony)	ꜱ	Stressed/Subject
ɪᴘꜰᴠ	Imperfective	Sd.	Sudanese
Kb.	Koalib	ꜱɢ	Singular
ʟ	Low	ᴜ	Unstressed
ʟꜱ	Low Set (Vowel Harmony)		

References

[A1] Anonym 1 (cca. 1950 ?), *Lessons in ŋirere*, Typed textbook designed to learn réré (**ŋireṛe**) in 30 lessons.

[A2] Anonym 2 (cca. 1960 ?), *Kawaleep Language, Yitam 1–7* [Livres 1 à 7], Abri ?, 108 pp.

[A3] Anonym 3 (cca. 1960 ?), *Ŋiṛaŋal ŋuluḏir zi kiṯam-na keni ṯikitaḏiza ṯuṛun* / Old Testament Stories in ŋirere, Abṛi (Nuba Mountains), 185 pp.

[A4] Anonym 4 (1996), *Yəwə na Nyaamin Nyəthi Kithla Kir 1996 Kandis-gi Kəthi Koaliib*, 24 pp. *without numbers.*

[A5] Anonym 5 (ca. 2000 ?), *Kiraw ŋa Kwurul gi*, Khartoum (?): no publisher mentioned, 2 p. [incomplete: the end of the story is missing].

[NT1] T̲ikitaḏiza ṯiaŋ (1967), Khartoum: The Bible Society of the Sudan.

[NT2] *Wa@d̲ wiyaŋ* (1993), Khartoum: The Bible Society in Sudan.

Abdalla, Jummize [Jummeiz] [CA1] & Komi, Abdalla (2000), *Yəwə na Nyaamin Nyəthi Kithilə Kir 2000*, Khartoum: Khartoum Workshop Programme. 8 pp. without any numbers.

Ewald, J. J. (1990). *Soldiers, Traders, and Slaves, State Formation and Economic Transformation in the Greater Nile Valley, 1700–1885*. Madison: University of Wisconsin Press.

Manfredi, S. (2010). *A Grammatical Description of Kordofanian Baggara Arabic* [PhD]. Naples: Università degli Studi di Napoli "L'Orientale".

Nunez, J. J. F. (2015). *L'alternance entre créole afro-portugais de Casamance, français et wolof au Sénégal. Une contribution trilingue à l'étude du contact de langues* [PhD]. Paris/Dakar: INALCO (Institut National des Langues et Civilisations Orientales)/UCAD (Université Cheikh Anta Diop).

Quint, N. (2006). *Phonologie de la langue koalibe, dialecte réré (Soudan)*. Paris: L'Harmattan. xvi + 238.

Quint, N. (2009). *The Phonology of Koalib*. Cologne/Köln: Rüdiger Köppe. Coll. Grammatical Analyses of African Languages. Vol. 36. xviii + 206.

Quint, N. (2010a). "La littérature orale en koalib (Sud-Kordofan)." *Études Littéraires Africaines (ELA)* (28): 45–57.

Quint, N. (2010b). "Benefactive and Malefactive Verb Extensions in the Koalib Verb System". In Seppo Kittilä & Fernando Zúñiga (eds.), *Benefactives and malefactives. Typological perspectives and case studies*. Amsterdam/Philadelphia: John Benjamins. 295–315. https://doi.org/10.1075/tsl.92.12qui

Quint, N. (2014). "Integration of borrowed nouns in Koalib, a noun class language". In Roger Blench & Thilo Schadeberg (eds.), *Nuba Mountain Language Studies*. Cologne/Köln: Rüdiger Koppe. 115–134.

Stevenson, R. C. (1984). *The Nuba People of Kordofan Province*. Khartoum: University of Khartoum.

Thomason, S. G. & Kaufman, T. (1991 [1988]). *Language Contact, Creolization and Genetic Linguistics*. Berkeley/Los Angeles/Oxford: University of California Press.

Winford, D. (2003). *An Introduction to Contact Linguistics*. Malden/Oxford: Blackwell.

Appendix 1. Phonological charts of Koalib vowels and consonants

Table 1. Koalib vowel chart (adapted from Quint 2009: p. 31; 2006: p. 31)

	Front	Central	Back
High	i		u
Mid	e	ɐ	o
Low	ɛ	a	ɔ

Table 2. Koalib consonant chart (adapted from Quint 2009: p. 49; 2006: p. 51)

Mode of articulation		Place of articulation					
		Labial	Dental	Retroflex	Palatal	Velar	Labio-velar
Obst.	Plain	b	d	ɖ	ɟ		
	Strong	pp [pp, p]	tt [tt, t]	ʈʈ [ʈʈ, ʈ]	cc [cc, c]	kk [kk, k]	kkw [kkwʷ, kʷ]
	Weak	p [p, f, v]	t [t, ð]	ʈ [ʈ]	c [ʃ, ʒ]	k [k, g]	kw [kʷ, gʷ]
	Pre-nasalized	mp [mb]	nt [nd]	nʈ [nɖ]	ɲc [ɲɟ]	ŋk [ŋg]	ŋkw [ŋgʷ]
NO	Nasal	m	n		ɲ	ŋ	ŋw [ŋʷ]
	Liquid		l / r	ɽ	j		w

Legend: Obst. = Obstruent; NO = Non-Obstruent. For the phonemes which have more than one allophone or whose transcription does not correspond to IPA, the attested phonetic realizations are given between square brackets. For more details, see Quint (2009; 2006).

Appendix 2. List of the 16 CÁCÈCÀ adverbs (and related forms) attested in my corpus of Arabicisms in Koalib

Type	Item	Meaning	Sd. Ar. root	Other borrowings from the same root in Koalib
(C)ÁC(C)EbCÀ	*ájèrà* [áɟèrà]	'hiring/renting'	[ʧr]	
	álèlà	'completing payment of dowry'	[ḥll]	
	állèmà	'studying/teaching'	[ʕlm]	*tàalêm* (n.) 'studies' < Sd. Ar. *taʕlíim*
	cápèrà [ʃávèrà]	'travelling'	[sfr]	*cèpèríè* [ʃèvèríè] (n.) 'journey' < Ar. Sd. *safaríyya(t)*; *cèprí* [ʃèvrí] (v.) < Ar. Sd. *sáfar*
	céjilè [ʃéɟìlè]	'recording'	[sʧl]	
	céŋkìrè [ʃéŋgìrè]	'rusting'	[ṣgr]	
	kêilè	'spending your time'	[gyl]	
	kérmìcè [kérmìʒè]	'wrinkling'	[krmʃ]	
	pékkìrè [fékkìrè]	'thinking'	[fkr]	
	táajèrà [táaɟèrà]	'making business'	[tʧr]	*táajèr* [táaɟèr] (n.) 'merchant' < Sd. Ar. *táajir*
	wénnìcè [wénnìʒè]	'chatting'	[wns]	
(C)ÁC(C)ÒCÀ	*ákkòmà*	'judging'	[ḥkm]	*àkkèmé* (v.) 'judge' < Sd. Ar. *ḥákam*; *ókkòm* (n.) 'judgement' < Sd. Ar. *ḥúkum*
	ápttòrà [áftòrà]	'having lunch'	[fṭr]	*pùttûr* [fùttûr] (n.) 'breakfast' < Sd. Ar. *faṭúur*
(C)ÁCCÀ	*ánycà* [áɲɟà]	'succeeding'	[nʧḥ]	
CÁCCÈ	*céllì* [ʃéllì]	'praying'	[ṣlj]	
CÓCCÀ	*dórrà*	'harming'	[ḍrr]	

Why linguistics needs an historically oriented Arabic linguistics

Jonathan Owens

Bayreuth University

One of the more recent, and certainly one of the most empirically well-founded accounts of language change is Labov's (2007) division between transmission and diffusion. The former results in gradual change via incrementation, the latter in larger and irregular change. This study examines the generality of this distinction, which was based on American vowel systems, against the rich history of Arabic. Five case studies are described in which it is shown that Arabic, like English, has striking instances of language stability across varieties as geo-diachronically separated as Emirati and Nigerian Arabic. By the same token, there are equally striking instances of widespread change due to contact. It is argued that in only one of these, Nubi (Creole Arabic), can diffusional changes be considered irregular, while in three others, Baghdadi Arabic (phonology), Uzbekistan or Central Asian Mixed Arabic (morphology and syntax) and Nigerian Arabic (semantics of idioms), the changes though of differing degrees of magnitude in their outcomes, cannot be said to be irregular. The study highlights two points: global criteria for defining the outcomes of transmission vs. diffusion are elusive, and Arabic, because of the ability to triangulate into different phases of its past, offers an unusually interesting insight into the workings of historical linguistic processes.

Keywords: transmission, diffusion, incrementation, language stability, idiomatic change, creoles, mixed languages, Baghdadi Arabic, Nubi, Central Asian, Mixed Arabic

1.   Transmission and diffusion[1]

An intriguing question in language history is why some things change, while others do not. This is one of the questions which goes to the heart of dialect differentiation

---

in Arabic. One framework for understanding this situation is Labov's distinction between diffusion and transmission. This defines general sociolinguistic parameters for predicting when and what sort of changes occur. The two terms define contrastive expectations and situations, as follows.

Diffusion

– Change implemented by adults
– Large scale and irregular change
– Change from above; importation from other systems

Transmission

– Change implemented by children
– Incremental
– Change from below; change from within the system
– Maintained over generations

Labov exemplifies these two aspects with the behavior of short vowel systems. Labov argues that the complex New York-city short vowel system was maintained relatively intact as migrants carried it to across the Midland region of the US, including Albany, Cincinnati and, exceptionally, New Orleans, all examples of change via transmission. Some elements of what Labov considers the original NYC system were lost, for instance a constraint against tense [a] in certain function words and in open syllables, but on the whole the basic structural elements were maintained. Diffusion is exemplified in developments in Illinois and Missouri where a 'New York' type system meets the Northern City shift system. Arising independently of the New-York City-type short vowel system, the Northern Cities shift, today centered around the southern Great Lakes, is simpler in that short vowels are not sensitive to phonological and morphological contextualization in the way the New York system is. In particular, whereas the short [a] in the NYC-system is sensitive to eight different factors, some phonological, some morphological (Labov 2007: pp. 354–355), the short [a] in the Northern Cities-system though itself part of a complex chain shift, is not subject to complex individual conditioning factors (2007: p. 372).

When Labov looked at the short vowel patterns in the corridor (I–55 corridor, named after the interstate highway) between Chicago and St. Louis, and in St. Louis itself, he found a system which was reminiscent of the Northern Cities shift, but still significantly different from it. Even though St. Louis is geographically in the Midland region, Labov describes it as an incipient Northern Cities system.

Labov is interested in establishing universal parameters of change, and in this he sees transmission as being the most important instrument in creating and maintaining diversity.

> The view I present here is that the primary source of diversity is the transmission (and incrementation) of change within the speech community, and that diffusion is a secondary process, of a very different character. (2007: p. 347)

On this basis, contact-induced change should be marked by a greater degree of irregularity than change via transmission.

However, Labov also allows that contact and diffusion may operate differently in other linguistic domains than it does in the low-level phonetic and phonological system which he studies in great detail.

2. Triangulation, chronological transparency

In this contribution I would like to explore the idea of transmission and diffusion as attested across different linguistic phenomena in regards to Arabic. As pointed out in the introduction to Owens (2013a: p. 9), Arabic provides an excellent laboratory for examining questions of historical linguistics precisely because, in contrast to the vast majority of languages and even language families, a large segment of Arabic is chronologically transparent, and what is important for comparative purposes, it is attested across many varieties. This latter property implies that the classic instrument of the comparative method, which can be termed "triangulation" can be employed in a wide range of circumstances, some of which will be exemplified here. Triangulation is simply deducing what might have happened to lead to situation A by comparing it with B and C. In the case of Labov's study, for instance, he observes the irregularity found in the St. Louis (as I will term the St. Louis-Chicago corridor) vowel system (A), and deduces on the basis of definable and regular properties of B (= Midland's short [a]) and C (= Northern Cities system) how A might have arisen (i.e. in this case, via diffusion).

To say Arabic is 'chronologically transparent' is a relative statement. It does not mean that Arabic is fully accessible and open to us beginning in the 7th century CE with the great Arabic-Islamic expansion. However, it does mean that in the classical tradition, particularly in the work of Sibawaih, we do have direct access to one form of Arabic as it was in the late 8th century. Moreover depending on how one judges the status of the *Qurʔaan* that was systematized by Ibn Mujaahid in the early 10th century, with the *Qurʔaan* we have an insight into Arabic even before this period. After this period our knowledge of the development of Arabic is indirect, for instance via the many 'Middle Arabic' texts (Blau 2002; Larcher 2001), until, roughly, the contemporary era when the many descriptions of Arabic dialects become available. This contemporary source, as will become clear in this article (Owens 1998, 2006/9, 2015b), in fact provides us the interesting possibility

of either reconstructing Arabic at different periods between the seventh century and today, or of allowing the conclusion that the contemporary dialects in fact cannot as a whole be derived from classical Arabic (Vollers 1906; de Landberg 1909), and therefore themselves need to be reconstructed minimally as parallel to Sibawaih's Arabic.

There is a second aspect of the linguistic history of Arabic which makes it a particularly compelling object to study using classic comparative methods, and that is, not surprisingly, in the course of its long, partly reconstructible history Arabic has developed in 'odd' and interesting ways in some cases, while in many others remaining very conservative. All judgments in this regard are based on linguistic parameters.

Note that linguistic parameters can be treated in two ways. In one, individual features are compared and a linguistic history of the feature is described. In the other, the parameters join to define varieties which either arise at points in the history of Arabic, or, as noted above, can only be traced back to an unanalyzable proto stage. Both perspectives will be treated in this article.

3.   A baseline measure of stability

Whereas historical linguistics is traditionally concerned with describing and explaining change, as soon as one has, as with Arabic, a roughly 1,400 year window of observation (ca. 650 – present) the question of language stability becomes prominent, particularly to the extent that the language intuitively is in fact stable. One need look no further than a typical verb paradigm – either perfect or imperfect – to conclude that Arabic in some key domains of grammar is extremely stable. Table 1 gives partial paradigms for the imperfect singular verb in a range of varieties of Arabic. The date at the bottom of the column states when the population of Arabic speakers first settled in the relevant region, or in the case of Classical Arabic, when a comprehensive version is first attested (i.e. with Sibawaih).

Table 1. Singular imperfect verb, Arabic varieties

Uzbekistan	Emirati	Cairene	Moroccan	Nigerian	Classical
1 *a-ktib*	*a-ktib*	*a-ktib*	*nə-ktəb*	*a-ktub*	*ʔa-ktub-u*
2M *ti-ktib*	*ti-ktib*	*ti-ktib*	*tə-ktəb*	*ta-ktub*	*ta-ktub-u*
2F *ti-ktib-iin*	*ti-ktib-iin*	*ti-ktib-i*	*tə-kətb-i*	*ta-ktub-i*	*ta-ktub-iin*
3M *yi-ktib*	*yi-ktib*	*yi-ktib*	*yə-ktəb*	*i-ktub*	*ya-ktub-u*
3F *ti-ktib*	*ti-ktib*	*ti-ktib*	*tə-ktəb*	*ta-ktub*	*ta-ktub-u*
710	pre-Islamic	640	700–1100	1400	800

It is clear that the paradigms are essentially identical across all varieties, with the same morphemes in the same forms in the same linear order. It is equally clear that the geographical expanse of the paradigms allows one to infer a commonality which goes back to a pre-diasporic origin, i.e. to the time frame which this article is concerned with. This quite obvious observation will lead to further interesting questions in historical linguistics, in particular why the paradigm should remain so stable over such a large separation of time and space. This is a question I will not go into in detail in this paper (see Owens and Dodsworth, ms.), beyond the summary observation that it is not so much paradigms, a linguistic abstraction, which are preserved as it is ways of using the paradigms in discourse. In particular, through their person affixes the paradigms track referents in discourse. The vast majority of Arabic verbal clauses lack an overt subject (Owens et al. 2009, 2013), but discourse identity of the subject is maintained in part through the verbal inflections. For instance, in the following extract from an Emirati text (see below), the change in subject to *ya-t*, feminine, though not marked by an overt subject, signals to the listener to find a feminine referent, which in fact they will have experienced four clauses previously in the feminine object suffix *-ik*. Since it is exceedingly rare for new subjects to be introduced with Ø, the listener will infer that it is the same feminine entity who is being talked about.

(1)   … (−4)

wu	yaa	iθ-θaani
and	came.M	DEF-other

And the other came

Ø rigad

Ø slept.M

And he slept

wu		Ø ya-t
and		Ø came-F

And she came

To bring this into an historical perspective, the hypothesis may be entertained that wherever one has the paradigm as in Table 1, the distribution of null and overt subjects will be roughly comparable, since the paradigms are not only paradigms, but also have a similar subject tracking function. This idea was tested (Owens and Dodsworth, ms.) using a parallel sample of oral discourse from contemporary Emirati (see Owens et al. 2013 for description of sample) and Nigerian Arabic (see Owens and Hassan, online resource). In all the approximately 25,000 word NA corpus has 2,520 clauses with verbal predicates while the Emirati (24,000 words) has 2,623 clauses. Table 2 shows how many verbs occur with null or overt subjects.

Table 2. Null/overt subjects in two Arabic dialect corpora

	Overt subject with verb	Null subject with verb
EmA	638	1985
NA	644	1876

As one can see by inspection, the distributions are virtually identical. What this suggests is that the paradigms in Table 1 are held in place, inter alia and perhaps for the most part, by their discourse function. It is not so much, and not only that the paradigms remain constant over 1,200 years, it is also the case that the underlying manner of discoursing, as measured by the expression of subject in discourse, has been maintained. Stability in Arabic can therefore be shown to extend well beyond the traditional historical linguistic domains of morphology and phonology.

This tells us that some things don't change. Transmission can be assumed as the operative concept in core domains of Arabic grammar. That aspects of the paradigm do change, for instance the generalization of the first person *n-* to the singular in North African (and a few others, Owens 2003) varieties can be understood as system-internal incrementation.

4. Four examples of linguistic change

Arabic is equally marked by change via diffusion. Some of the four cases discussed in this section are characterized by striking and unmistakable instances of diffusion, but others are more subtle and need to be argued for in some detail. I begin with one of the latter. I would add that each of the following four case studies are synopses of studies which are described more thoroughly elsewhere, where they can be consulted for detail.

4.1 Syllabification

The first example comes from one of the traditional domains of comparative grammar, phonology. In (2) are three paradigms for the perfect verb, two Arabic, one Biblical Aramaic. Clearly, the surface form of the Baghdadi Arabic paradigm is in terms of syllable structure more similar to Biblical Aramaic than it is to its sister Cairene Arabic. Cairene, essentially, maintains its stem vowels as they are, whereas Baghdadi and Biblical Aramaic exhibit a complicated alternation of vowels. However, there is a complex but coherent phonological rule which accounts for the Baghdadi and Biblical Aramaic similarities (Aramaic examples from Rosenthal 1961: p. 60ff.).

(2)   'write', perfect verb

Cairene Arabic		Baghdadi Arabic		Biblical Aramaic	
SG	PL	SG	PL	SG	PL
katab-t	*katab-na*	1 *ktab-it*	*ktab-na*	*kitb̠-et̠*	*k̠t̠ab̠-na*
katab-t	*katab-tu*	2M *ktab-it*	*ktab-tu*	*k̠t̠ab̠-t*	*k̠t̠ab̠-tu*
katab-ti		2F *ktab-ti*		*k̠t̠ab̠-ti*	
katab	*katab-u*	3M *ktab*	*kitb-aw*	*k̠t̠ab̠*	*kitb̠-u*
katab-it		3F *kitb-at*		*kit̠b̠-at̠*	

The following constraints account for both Baghdadi and Biblical Aramaic, which will be presented in very abbreviated form from Owens (2017).

(3)   Constraint and repair system (C-R)
    a.   Short vowels do not stand in open syllables. A short vowel in an open syllable is deleted.
    b.   Inappropriate consonantal sequences which arise via (3a) are broken up by insertion of an epenthetic vowel between C2 and C3.
    c.   Insertion occurs in the following contexts (counting from the word end).
       Sometimes in C3_C2C1.
       Always in C4C3_C2C1.
       Always in C3_C2#1.

This is termed a constraint and repair system (or schema) because its basis is two constraints, one which disallows a (unstressed) vowel in an open syllable, but also disallows a sequence of three consonants, and a repair mechanism, namely the insertion of an epenthetic vowel (3b, c) when a CCC (or CC#, #CC) sequence arises, either via suffixation or via the deletion of a vowel in an open syllable.

(4)   Baghdadi Arabic
    a.   *kítab-at* 'she wrote' > *kítb-at*   (via 3a) no V in CV-CV
    b.   *kitáb-t#* 'I wrote' > *ktáb-it#* (or *kitáb-it*)[2]
                                via 3a, c, condition 1, no CC#, so CəC#

(4a) finds the /a/ in boldface in an open syllable, against (3a). It needs to be added here that stress protects against deletion, so the stressed initial 'i' is protected. In

(4b) the final sequence b-t# has the equivalent of a three-consonantal sequence (# = final C), so via (3b), (c) an epenthetic vowel gets inserted in the b_t# sequence. Note that the epenthetic vowel is 'invisible' to the Baghdadi stem. After its insertion the stem vowel /a/ in *ktab* is in an open syllable, but this does not trigger (3a) since the following vowel is epenthetic.

The same set of constraints accounts for the Biblical Aramaic paradigm.

(5)   Biblical Aramaic (using the traditional underlining for a fricative, <u>t</u> = θ)
    a.   *kitab-at > ki<u>t</u>b-a<u>t</u>* via (3a), no V in an open syllable CV-CV
    b.   *k<u>t</u>a<u>b</u>-ti* via (3a) this time acting on first stem vowel

There is one difference between Baghdadi and Biblical Aramaic as shown in (6)

(6)   *ki<u>t</u>b-e<u>t</u>* 'I wrote'

This is, however, not a difference in the syllable structure constraints but rather in the status of the vowel of the 1SG suffix. Historically, the final /e/ is supplied by the 'same' constraint 3b, c which gives the vowel in the 1SG suffix in Baghdadi Arabic. In contrast to Baghdadi, the epenthetic vowel on the 1SG person suffix does count as a systematic vowel (Segert 1997: p. 122), which induces deletion of the preceding vowel in an open syllable.

Allowing for a number of individual differences of interpretation, the constraint and repair schema is attested throughout the classical Aramaic languages where short vowels are attested in script (Syriac and Mandaic, Nöldeke 1898: p. 37; Malone 1971, 1997; Muraoka 1997: pp. 10, 35, 143; Voigt 2007: p. 162). It is also noteworthy that the Baghdadi paradigms, and, presumably, the underlying rules, are also attested in NW Syrian Arabic (Behnstedt 1997: p. 142ff.).[3]

The identity of the Baghdadi Arabic and Biblical Aramaic engenders three basic issues in historical linguistics which can be dealt with in summary fashion here. A starting point, (see Moulton 1954: p. 38) on which other questions depend, is whether the Baghdadi and Aramaic paradigms are the result of independent parallel development or not.

There are no hard and fast criteria for discerning parallel independent development vs. shared development, but it is clear that the simpler the phenomenon the more difficult it is to distinguish the two, and by the same token, the more complicated, the easier it is. For instance, both Bagirmi Arabic in eastern Nigeria and Uzbekistan Arabic share the unconditioned shift of *θ > s, *sagiil*/*saqiil* 'heavy' < *θaqiil). θ > s is attested elsewhere in Semitic (e.g. Akkadian), and is found

---

3.   Ignoring for purposes of this abbreviated exposition the issue that many Arabic dialects have elements of the constraint and repair schema to one degree or another, without being identical to the Baghdadi Arabic. This issue is discussed in greater detail in Owens (2017).

in L2 varieties (e.g. German L2 of English).[4] It is probably impossible to decide in this instance whether independent or shared innovation is in play. In the present case, however, one is confronted with a more complicated set of procedures. Still, they break down into basically three. Given these, one might argue that once the general idea enters the language (i.e. acquires speakers), the constraint and repair schema of (3/7) will emerge independently.

(7) The constraint and repair system (C-R system)
    a.   No short vowels in unstressed open syllables
    b.   No sequences of 3 consonants
    c.   In case (b) arises, insert vowel between consonants

Once these conditions enter a system, it could be argued, then there is a high probability that eventually identical paradigms will be reached along independent paths of development. This is what lies behind for the striking identity of the Baghdadi Arabic and Biblical Aramaic perfect verb paradigm illustrated in (2).

This explanation can be countered in three ways. First, dialects which have (7a–c), Baghdadi Arabic and NW Syrian Arabic, for instance–are also those where historically intensive Aramaic-Arabic contact and bilingualism is highly plausible. Any argument pertaining to diffusion needs a plausible demographic justification of contact.

Secondly, as Retsö (2000) shows, there are a number of significant features beyond these which point in the same direction of change within Arabic via early contact with Aramaic.

In this regard, the assumption of a common C-R schema leads to obvious parallels between Aramaic and Arabic forms. A case in point is illustrated in (6) above. The interpretation of the Biblical Aramaic *-et* as containing an epenthetic vowel (Segert 1997: p. 122) follows from its parallels with Baghdadi Arabic, and fits in with the behavior of epenthetic vowels in Aramaic.

Thirdly, the complex of constraints and processes in (7) taken as a whole is hardly found elsewhere among world languages, even if the general pattern falls within what are termed 'conspiracies' in phonology. In a phonological conspiracy, different rules or constraints work towards a common end, for instance, towards preventing a final CC# cluster or medial CCC cluster. In certain respects, rules very similar to those in Aramaic/Arabic considered here have been described elsewhere, for instance in the Californian Native American language Yokuts (Kisseberth 2011 for summary). Even within Afro-Asiatic it is easy to find two of the three steps of (3/7). In Oromo, for instance, CC-C requires repair to either C**a**CC or

---

4.   Recall the Berlitz advertisement for English L2 in Germany, 'we are sinking!', coast guardsman monitoring the distress call pauses, then replies, 'what are you sinking?' (i.e. thinking).

CC**i**-C, depending on the nature of the consonants involves (e.g. *kofl-te → kofal-te* 'I laughed'). Similarly, more complicated three-step 'bleeding and feeding' routines can also be found. Paradis (1988: pp. 77–79) for instance describes a three-step process in Fula of the type $CC_{continuant} > C_{continuant} > CC_{stop}$ (e.g. *ss > s > cc*) where the intermediate step, comparable to the CCC sequences in (10), (13e), feeds into the final step.

What is unusual, however, is not the C-R system in the abstract, but rather the syllabic domain it is applied to, its instantiation via epenthetic vowel insertion, and its regularity, with step (3a/7a) feeding into (3b, c/7b, c). All of these elements are shared between Arabic and Aramaic.

Assuming then that the identities in the two paradigms in (2) have a single origin, the next question to ask is whether they indicate a common inheritance or borrowing from one language to the other. Given that there are Arabic varieties, for instance Cairene in (2), which have only limited elements of the C-R schema, the most likely explanation is one of borrowing. The historical record lends plausibility to this. For instance, Lapiński summarizing the relations between Aramaic and North Arabian speakers writes,

> … the global history of these Aramaeans in the 8th–7th centuries BC can hardly be separated from the history of the North-Arabian tribes living in the same regions and called "Aramaeans" in Assyrian sources that barely and only exceptionally distinguish the two groups. (2000: p. 485)

Similarly Retsö (2003) amply documents extensive contact between the two groups well into the BCE era and continuing into the Islamic era. Given that in this period Aramaic, beginning in the Achaemenid era ca. 600 BCE was the lingua franca in the region it stands to reason that borrowing would have gone from Aramaic to Arabic. What remains to be worked out is when the contact would have taken place, already in the Middle Aramaic era (450 BCE–350 CE), or perhaps later.

Given that the systematic identity in this case involves a complex structure, clearly intensive contact would have been involved. This could equally have been via borrowing or shift. If borrowing, L1 Arabs would have been highly fluent in the Aramaic lingua franca which dominated the region for centuries and shifted Aramaic syllable structure into Arabic. The multilingual environment would have been amply present for this to have happened. If via shift, L1 Aramaic speakers would have brought their Aramaic syllable structure into Arabic (van Coetsam/ Winford's "imposition"). This presumably would have occurred at the earliest at the end of the period given above, when Aramaic began giving way to Arabic as the lingua franca.

In either case the specificity of the complex C-R system makes it highly likely that it arose only once, probably in Aramaic, and made its way via diffusion into some ancestral Arabic dialects.

4.2   Semantic change

Beginning in this section I would like to look at three domains of change which fall outside the traditional purview of language change studies, either because they have in fact received relatively less attention in Linguistics (4.2), or because by their very nature they are unusual (4.3, 4.4).

For the next case study we move back to NE Nigeria. As already seen in Section 2 above, NA displays not only classic paradigmatic features of Arabic, but also distributes them in discourse in a manner which is identical to at least one other variety separated from it by approximately 1,200 years. NA, however, does not uniformly display features from other varieties of Arabic. In fact in one respect, NA is closer to Kanuri and other languages of the Lake Chad area than it is to 'Arabic'. This component concerns idiomaticity.[5] The contrast can be indicated by the following 'minimal' pair between NA on the one hand and EA/STA Egyptian Arabic/ southern Tunisian Arabic on the other, invoking NA and STA in this comparison since it is known from historical (written) sources that both NA and STA migrated out of Upper Egypt to their current locations. The ancestral STA population (Banu Hilal) left Upper Egypt around 1150, while ancestral NA speakers had reached the Lake Chad area from Upper Egypt by 1390. All three dialects therefore share a common immediate origin.

(8)	minimally contrastive idioms	NA	EA	STA
	{[ṛaas-PSSR yaabis[6]] head dry}	dumb	stubborn	stubborn/ narrow minded
	{[gaḷb baarid] heart cold}	easy going	indifferent, lazy	indifferent, lazy

In NA a {dry head} is a dumb person, whereas in Egyptian and southern Tunisian Arabic it is a stubborn person or narrow-minded person. In NA a {cold heart} is an easy going person whereas in EA and STA it is a lazy or indifferent person. The same lexical collocates have contrastive meanings.

---

5.   I will not attempt to characterize what I mean by idiomaticity here; see Owens (2015a) and ms. for detailed discussion.

6.   *našfa* in EA, *naašif* in STA, *yaabis* being 'dry' in NA, which does not have the lexeme *naašif*. I enclose idioms in curly brackets, "{}".

The examples in (8) are typical of a wide range of idioms in that EA and STA show a very high degree of agreement in their idiomaticity, as opposed to NA. The following examples, from Benmimoun et al. (2017), further illustrate this point, using examples from the rich idiomaticity associated with word for 'head', *ṛaas*. EA and STA share both the collocates and the idiomatic meanings the collocates engender.

> (9)   'one unit of'
>   STA
>   ṛaas X = one unit of, {*ṛaaṣ biṣal* 'head of an onion = an onion}, {*ṛaaṣ ɣanam* head of a sheep = 'a/one sheep'}.
>   EA
>   ṛaas X = one individual type {*ṛaas baṣal* 'head of an onion = 'an onion, *ṛaas ɣanam* = head of a sheep = 'a/one sheep'}

> (10) 'be dizzy'
>   STA
>   {*ṛaaṣi yduuṛ* my head is spinning = I am dizzy}.
>   EA
>   {*raas-u daar-it* his head spun "He got dizzy"}

> (11) 'humiliate'
>   STA
>   {*ḥaṭṭ-l-a ṛaaṣ-a fi t-ṭraab* he put his head in the ground = he humiliated him}.
>   EA
>   {*ḥaṭṭ raas-PSSR fi t-turaab* [put head in the dust] = humiliate}.

In this case these collocations in NA are either meaningless (*ṛaas baṣal* for instance,) or produce only a literal meaning (*ṛaas qanam* = 'a literal head of a sheep').

By the same token, Nigerian Arabic idioms are typically those which are equally nonsensical, or literal, to non-NA speakers. As shown in Owens (2014b), NA idioms typically are calques from Kanuri. The following three examples based on the idiomatic keyword *ṛaas* 'head' on the one hand show the collocational identity between Kanuri and NA, and on the other indicate how different the NA collocations are from, for instance, Egyptian Arabic.

(12) Kanuri – NA idiomatic identities

Kanuri	NA	common meaning
(a)	ṛaas X = bundle	
kəla kajim-be	*ṛaas al-gašš*	
head grass-of	head DEF-grass	bundle of grass
(b)	ṛaas X = tassel	
kəla argəm-be su-tulug-əna	*ṛaas al-qalla marag*	
head corn-of s-come	head DEF grain come	'tassels have appeared'
out-PST perfect	out	
(c)	ṛaas X = roof	
kəla fato-be	*ṛaas al-beet*	
head house-of	head DEF-house	roof

In both Kanuri and in NA a bundle of grass of the type which is tied and can be carried on the head is a 'head of grass', using the NA words for 'head' = *ṛaas* and 'grass' = *gašš*, the Kanuri using Kanuri 'head' = *kəla* and 'grass' = *kajim*. In both there is a possessive relationship between the nouns. In EA the collocation *ṛaas hašiiš* or *ṛaas al-ʔamħ* is odd. In (12b), the expression for 'tassels' coming out both in Kanuri and in NA uses the lexemes for 'head' and 'come out'. Something in the range of 70–80% of all NA idioms collected from a corpus of 400,000 words of spoken text (see Owens and Hassan) are based on lexical collocations which are calques from Kanuri.[7] For some basic lexemes these idiomatic collocations are their 'normal' state. For instance, in the 400,000 word corpus, *gaḷb* 'heart' appears in 101 tokens, all of them idiomatic, none of them the literal heart (see Owens 2015a).

Clearly NA idiomaticity is the result of diffusion. It is also pervasive. As noted and exemplified above, a comparison between NA, EA and STA revealed very few commonalities in idiomaticity between NA and EA/STA (Ritt Benmimoun et al. 2017).

There are marginal instances of syntactic borrowing in NA, in particular in the NP. Virtually all noun modifiers are post-noun, including the quantifiers (numerals, 'all'). Whereas in most varieties of Arabic the quantifer for 'all' is pre-head, in NA it can only occur post-head.

---

7. Indeed, as shown in Owens (1996), the NA-like idiomaticity is spread throughout the languages of the Lake Chad region (e.g. Bura, Kotoko, Fulfulde of N. Cameroon), this reflecting the long period of Kanuri hegemony in the region. For the non-overlapping idioms, most are equally different with EA, though their origin has not been identified. Some are probably internal creations.

Areally-defined calquing has been noted elsewhere, e.g. for Mesoamerica Kaufman et al. (1986: pp. 553–554), though this phenomenon is not usually defined in relation to idiomaticity,

(13)   a.   *šuf-t   kull in-naas*
          saw-I all   DEF-people (EA)
     b.   *šif-t   an-naas   čattu-hum*
          saw-I   DEF-people all-their
          'I saw all the people' (NA).

However, as can be seen in comparing it to the striking stability of NA in the domains of verbal morphology and discourse attributes of subjects demonstrated in Section 2 above, the only fundamental way contact has influenced NA is limited to the domain of figurative language (see also Owens 1994, 1996 to this point).

4.3   Mixed Arabic: Uzbekistan and Afghanistan

Among the varieties which Arabic has acquired in its variegated history is what appears to be a mixed language, spoken by a small community in Bokhara, Uzbekistan.[8] In the eighteenth century speakers of this community migrated to northern Afghanistan as well (Kieffer 2000: p. 183).

    Unfortunately there are no comprehensive descriptions of this variety, though from what we have it is clear that it has 'classic' characteristics of Arabic, but also elements from the co-territorial languages it has been in contact with since 710, when Arabs first moved into the region, as well as striking, original innovations. For instance, as can be seen in Table 1 it has an inherited verb paradigm which is unmistakeably Arabic and it equally has the familiar Arabic object pronouns (see for instance (14) *-hen*). By the same token, its word order is clearly that of co-territorial Tajik/Dari (= Farsi varieties) and Uzbek being SOV.

(14)   a.   boqiir *hat kuusa ḥallaa-hen*
          cows   this youth freed-them.F
          'The boy let the cows loose'   (Versteegh 1984–6: p. 448)
     b.   baqara *m-a-rʕee*
          cows   IND-I-graze
          'I graze cattle'.   (Kieffer 2000: p. 192)

Beyond this it has acquired, via adstratal languages, many discourse particles (*-mi* 'Q marker < Uzbek via Tajik, *-u* 'and', *ki* 'complementizer', < Tajik, Fischer

---

8.   There is also reported to be/have been a group of Arabic speakers in two villages in Kashka-Darya province in south-central Uzbekistan. According to information supplied to me by Dr. Talal Aljassar, who visited one of the villages, Djeinau in 2015, only four or five very old speakers of this dialect are to be found.

    While I believe the term "Central Asian mixed Arabic" is a more accurate designation than "Uzbekistan Arabic", the latter will be more familiar to Arabicists.

1961: p. 260–262), like Tajik/Dari and Uzbek has what are former demonstratives as the third person independent pronouns, e.g. *haad/haaz* 'he' (other dialects = 'this.M'), *hai/haadi* 'she' (other dialects = this.F; Zimmermann 2002: p. 31), does not have a definite article (*il-* or *al-* for instance) to signal discourse givenness or genericness (see (14)), and appears to have a reduced set of broken plurals (Fischer 1961: p. 243; Kieffer 2000: p. 185). Like Tajik it makes extensive use of a light verb system consisting of verbal noun + *sawwa* 'make, do', *šoɣol sawee-na* 'we made work = we worked, cf. Persian *kar kardan* 'work make = work'). Except for codeswitching, Arabic dialects do not have this construction. In the verbal system it appears to have re-functionalized the active participle into a perfective verb (whose function needs more detailed study) in such a way that the participle is fully inflected for subject markers, parallel to the original finite verb. How it has done this is as follows. It keeps the gender marking for third person forms, thereby maintaining the gender distinction as in (15). Since these forms do not have *-in* (see below), a contrast is thereby created between third person and the other persons.

(15) *qaʕd-a*
      sitting-F
      'she has sat'
      qaaʕid
      sitting
      'he has sat'.

For the first and second persons it uses what is originally the active participle, plus the auto-morphological 'intrusive -n', followed by the obligatory subject person marker, which is nothing more than the 'former' object pronoun which is obligatory after the -n. (Owens 2013b). Zimmermann (2002: pp. 45–46, 2009) gives the following examples.

(16) *qaʕd-in-ki*
      Sitting-n-you.F
      'you.F have sat'
      zorb-in-kii-hum
      hit-n–2.F.SG-them.M
      'you.F.SG have hit them.M.

Clearly the variety has undergone considerable changes, whose components in contrast to the creole case discussed in the next section, still have a transparent historical origin: finite verb inherited from Arabic, word order and other fundamental grammatical features from co-territorial Tajik/Dari and Uzbek, an innovative verb

system, foreign to Arabic in its systematic values,[9] but transparently derived from the active participle.[10]

Uzbekistan Arabic shows classic features of a mixed language (e.g. Michif, see Bakker 1997) in that its interesting complexities are transparently derivable from a fusion of different languages. Since we know that historically it is Arabs who moved into Uzbekistan after 710, the differences clearly diffused from Tajik/Uzbek into Arabic.

4.4   The limits of Arabic: Creole Arabic

The last position we come to on our scale is Creole Arabic. This is a variety which arose in the second half of the nineteenth century in, roughly, the area which today is the new sovereign state of South Sudan, and established itself as a native language in this region and in East Africa (today Uganda, Kenya) in the early twentieth century (Owens 1997; Wellens 2005; Miller for Juba Arabic). In East Africa this variety is known as Nubi, a term which I will use for it here. That this is derived from Arabic, yet not itself a variety of Arabic is intuitively clear when the paradigm from Table 1 is repeated with Nubi values.

(17)   Nubi imperfect verb paradigm
sɢ 'katifu 'write'
1   *ana gi-katifu*
2   *ita gi-katifu*
3   *uwo gi-katifu*

---

9.   See discussion in Retsö (1988). Retsö emphasizes the role of the co-territorial languages in shaping the overall verbal system of the variety.

10.   Indeed, the fully person-inflected AP arguably makes the verbal system more transparent and regular than the inherited Arabic one, where the AP, though part of the verbal system, is morphologically adjectival.

Windfuhr (2005: p. 118) draws a parallel between this Uzbekistan Arabic neologism and Sulaimani Kurdish, which also has an inflected participle in perfective function. Unless a direct Kurdish-Uzbekistan link can be drawn, I think a better place for an immediate source of the new inflected perfective is in Aramaic, which in 710 is known to have provided a significant substrate in the early diasporic populations (see 4.1 above). Aramaic, as is well known, has developed a verbal system based on the active participle, going back at least to Biblical Aramaic. Macuch (1982: pp. 118, 204) notes that the (formally) second person object suffixes in Samaritan Aramaic serve as subject markers on active participles, providing a more direct link (Samaritan attested ca. 100–700 CE). The parallel with Kurdish would attest to a broader sprachbund phenomenon. For further treatment of Aramaic-Arabic diffusion, see Contini (1999).

In contrast to Arabic, an overt subject is required, there is no internal ablaut in the verb, only a single segmental stem form, and the imperfectivity is shown by an explicit prefix, *gi-*. Each of the morphemes is etymologically Arabic – *gi-* derives from *gaaʔid*, a marker of progressive in the Arabic of Chad and Sudan Arabic (as well as occasionally other dialects) – but clearly a massive restructuring has occurred. Often Nubi is simpler than corresponding structures in Arabic, as here the imperfect verb lacking the complex morphology of Arabic. However, this is not universally so. In some cases Nubi displays unpredictable or even unique complications (Owens 2014a). Linguistically speaking the transition from Arabic to Nubi can be seen as using Arabic morphemes in new structures which often bear only limited resemblance to the lexifier language.

The reason Creole Arabic came into existence is transparent, and repeats a story which lies at the origin of most creole languages. A heterogeneous, multilingual community from the southern Sudan living in the presence of a relatively small dominant group speaking Arabic (northern Sudanese, Egyptian officers and those from other Ottoman-controlled Arab countries) had within the period of two generations to 'consensualize' a lingua franca. The dominant group was too small, and probably was not particularly motivated, to impose "normal" Arabic on the dominated linguistically heterogeneous groups, with the result that Arabic was used as a basic morphemic building block of the lingua franca, though in ways which produced a new grammar, stabilizing in a new language (Owens 1997).

It is hardly accidental that creolization within the broad domain of language contact has engendered a niche of acquisition theories distinctive from normal language transmission, whether L1 or L2. Creoles, as Bickerton (1981) rightly emphasizes, do not have a 'normal' history of language origin. There are many perspectives on this point, but here I will introduce one which has been appealed to understand the relation between Nubi and other varieties of Arabic. The basis of this perspective is the idea that Nubi is not particularly different from other Arabic dialects. Versteegh (1984) originated this idea in proposing that Arabic dialects underwent a stage of simplification akin to pidginization, which explains a purported fundamental difference between Classical Arabic and the dialects. A basic problem with this idea is that it is questionable whether Classical Arabic is in fact so structurally different from the dialects that one need even appeal to the idea of pidginization in them. Leaving aside many individual background issues, for instance what we know of the early Arabic expansion did not produce anything like the social disjunctions which are documented in the late nineteenth century southern Sudan, there exists a fundamental comparative linguistic issue, raised by Bergsträsser (1928: p. 156). Bergsträsser namely suggested that there is a set of innovations among the Arabic dialects as a collective which justifies assuming a fundamental historical linguistic break between Classical Arabic and dialects. This

alone is highly problematic (Owens 2006/9),[11] and hence the whole assumption that there is a case to be answered for in regards to the simplification of the dialects can be questioned.

Whereas Versteegh seeks to assimilate Nubi to the Arabic dialects via creation of an historical analogy (what happened to the dialects is like what happened to Nubi), McWhorter (2007: Chapter 7) takes an opposite approach, namely by creating a cline of simplification beginning with dialects and ending in Nubi. McWhorter's approach, in fact, is somewhat surprising, since elsewhere he has been a strong advocate of 'Creole exceptionalism' (e.g. 1998, 2002, 2012). Creoles are what they are, in part because they have gone through a pidginization stage. McWhorter, in contrast to Versteegh, does not envisage such a stage in the Arabic dialects.

Rather than try to accommodate the dialects to Nubi, or Nubi to the dialects, it is better to see that at this juncture a point of qualitative difference has been reached. The conditions for language contact, for language learning, for the transfer of structures from L1 Arabic to a new group of L2 learners was present in the nineteenth century southern Sudan, but the end result escapes the, by way of comparison, 'simple' outcomes discussed in 4.1–4.3 above. Nigerian Arabic, despite its strikingly new (for Arabic) idiomaticity is still Arabic; if Baghdadi Arabic owes fundamental aspects of its syllable structure to Aramaic, it is still Arabic. Nubi, however, is not, a conclusion appropriately implied in the very concept of 'creole'.[12]

5.   Transmission and diffusion

To conclude, I return to the general question of diffusion and transmission as characterized by Labov (Section 1). Taken at face value the range of studies here indicate that a simple correlation between transmission/diffusion and how the effects of these manifest themselves is difficult. They certainly don't support the idea that diffusion is always marked by irregular change, transmission by regularity, probably the most interesting result from Labov's detailed case study. Idiomaticity in Nigerian Arabic is, intuitively, completely regular. Its irregularity resides merely in the fact that it has a regularity understandable relative to Kanuri, not other varieties

---

11. Bergsträsser himself never specified what common innovations bound together dialects against Classical Arabic.

12. Though even here some scholars might take issue. Mufwene (2009) and (2010), for instance, argues that creoles can be accounted for by a natural evolution from their various inputs, substratal and superstratal.

of Arabic.[13] In some cases they do support Labov's idea that diffusion causes large scale changes, Central Asian mixed Arabic and the massive re-idiomatization of Nigerian Arabic being two examples. On the other hand, one would probably not think of the Aramaic-Baghdadi Arabic diffusion as being such. However, there are further perspectives to be added and caveats considered in using Labov's study as a basis of comparison.

First it can be argued that a simple juxtaposition of outcomes mixes apples and oranges in different ways. One is the factor of time scale. Labov reconstructs the irregular St. Louis short vowel system to at least 1860 (2007: p. 382), lasting into the present. Though not an insignificant time frame, in the current study one is dealing in some cases, as with Baghdadi Arabic, with a time scale which in the final analysis can only be guessed at. Something happened, it can be plausibly argued, but when it happened at very best might be postulated as sometime between 400 BCE and 800 CE, the point at which Arabic displaced Aramaic as lingua franca in the region. Even if during the initial contact phase ancestral Baghdadi Arabic was less regular than today, such variation would have long since been regularized away. Similarly, Uzbekistan Arabic may have begun taking its current shape as early as 710 and Nigerian Arabic its special characteristics in the late fourteenth century. It stands to reason that in their incipient phases idiomaticity in Nigerian Arabic or word order in Uzbekistan Arabic would have been marked by variation between inherited Arabic and coterritorial language influences, but such variation lies too far in the past for detailed reconstruction. One is left only with the outcomes.

By the same token, however, the Nubi case shows that chronology itself is not a variable with determining force. It is known that Nubi became a language between 1851–1888 (Owens 2014a), i.e. stabilized within a very short period of time, within an era contained in the postulated period of St. Louis vowel variation.[14] Labov himself remarks that relatively large changes in a vowel system can occur in just three generations (2007: p. 379). It is not time in the abstract, but rather the social factors and degree and type of coherency among the community of speakers within a given chronological window which are crucial. This point is brought out nicely in Table 1: something must keep the paradigms in place across time and geography; a stable social factor is a key element Labov sees in the "normal" transmission of the

---

13. Assuming formalized semantic representations of idiomaticity in NA and in Kanuri would be essentially identical in structure. Such a representation is developed for NA (Owens 2015a) though linguistic 'proof' of correspondences must await a full scale analysis of Kanuri idiomaticity.

14. This is slightly less than a typical three generations which has been argued to produce a koine (Operstein 2014: p. 5). Note, however, that Labov sees diffusional irregularity in St. Louis lasting for well over 100 years (probably the upper limit of three generations); similarly, NA in urban Maiduguri shows no sign of koinizing after three generations.

NYC-system and the Northern Cities system within their "heartlands". The point is reinforced in the observation from Section 3 that it is not only paradigms which are stable but also the way of deploying them in discourse. Discourse implies shared norms of interaction.[15] Chronological time therefore is not necessarily a factor relevant to the linguistic outcomes described here.

Secondly, Labov's analysis of diffusion is in part a product of his methodology. Being based on variationist data from nearly 100 speakers he is able to statistically define the lack of homogeneity in St. Louis against that which is found in the Northern Cities and New York City-type data sets. To the extent that this is the case, one will best be able to compare his analysis to data sets which have variationist data, which by and large limits a comparison to data derived from contemporary situations.

It stands to reason, as noted in the previous point, that variationist methodology will more likely uncover irregularity than data reconstructing events which happened a millennium or so ago. However, as the Northern Cities and NYC-type data sets show, as well as the Emirati-NA discourse data, variation does not mean irregularity. To the contrary, in demographic terms it is the less regular St. Louis short vowel system which represents a smaller population than either the Northern Cities or NYC-type. Systems do regularize and if they change incrementally they change regularly. If the St. Louis system can be understood as the product of contact between two stable systems, it is a product which itself has been in a sort of equilibrium, if Labov's pre-1860 date for the start of the pattern is correct. If language change can be understood teleologically, St. Louis short vowels will eventually look like Detroit's. If it isn't, however, one will need to continue tracking its development to determine when, if ever, greater regularity sets in, and what form it takes. The situation is analogous to Arabic in Maiduguri (Owens 1998). Arabic is native to this relatively young (founded 1906) city in NE Nigeria for three generations. The rural inputs into Maiduguri are quite homogeneous. The urban outcome is irregular in ways reminiscent of the St. Louis vowel system. While the social status of Arabic is different in one important way, namely that Arabic in Maiduguri is a minority language, at this point there is no indication that Arabic is converging towards a common norm.

A crucial third factor involves the nature of the linguistic components in contact. To begin with, a vowel system is defined inter alia by physical values. By definition values are discrete and irregularity measurable against a comparative baseline, such as in this case the Northern Cities and the NYC systems represent. By way of contrast, there are no discrete semantic values, no physical baseline against which

---

15.   Indeed, the status of the speech act as situational, involving speaker and addressee, is one of the linguistic variables used in the Emirati and NA data.

idiomaticity can be measured.[16] Even quantified ordinal data as with the discourse values in Section 3, is defined relatively, one data set against another. The discourse categories themselves, moreover, are sometimes open to different interpretations. Looking at word order, should, one assume that a wholly 'harmonic' (Greenberg, Hawkins) word order (modifier – head or head – modifier), is an ideal and anything deviating from it an irregularity? On this basis the postpositions in Uzbekistan Arabic are analogous to the St. Louis vowel system, introducing a category otherwise absent in Arabic that makes for an overall greater structural irregularity. However, it was equally seen in (13) that NA in a sense regularizes NP elements, for instance the expression for 'all' in aligning it in a head – modifier order, in contrast to virtually all other varieties of Arabic where *kull* 'all' has a pre-head option. This post-head position probably is due to the influence of Kanuri and other African languages in the region, where the N is strictly NP initial. In this case diffusion increases typological regularity, in the sense used here (see Thomason and Kaufmann 1988: pp. 14–34 for discussion of related points).

Directly relevant in this domain, the treatment of diffusion has been presented across different domains, phonological (4.1), semantic (4.2), and what might be termed general linguistic systems (4.3) and (4.4). It is well known that different linguistic domains may tendentially[17] have different historical trajectories. Semantic systems are more transportable than word order changes and these in turn more than morphological ones. The fixedness of the Uzbekistan and NA verb follows this trend, and the fact that NA is strongly influenced by Kanuri in the semantic realm, but not syntactic – Kanuri is a fairly harmonic SOV modifier – head language – fits into this tendency as well. Words are more open to reinterpretation and reorganization than are their constituent morphemes. Idiomatic change is not only a change in meaning, but also a change in the collocational frequencies of individual lexemes; word order works on word categories.

Even in phonology, however, diffusion does not imply irregularity. This is the interpretation of the origin of the Baghdadi Arabic syllable structure. Parallel to Labov's study, in the domain of vowels, Kerswill et al. (2008) document the reversal of diphthong patterns in favour of shorter diphthongs or even monophthongs with a geographical center in inner London and spreading via diffusion outwards. While they do not attempt to analyse the data in terms of irregularity, it is hard to see in

---

16. This issue requires separate treatment. While meanings are unquantifiable (in my view), the fixed collocations which constitute idioms are.

17. Linguists such as Campbell (1993) and Thomason and Kaufman (1988) caution that there are few if any absolute constraints on what can diffuse.

the generalized summaries of the two systems (2008: pp. 482, 484) how the new diffusing pattern can be seen as less regular than the one it is displacing.[18]

Dialect contact vs. language contact. A final variable is whether the contact is intra- (dialectal) or inter-language. In all of the Arabic cases I have only considered inter-language contact. The potential for radical change via contact is greater for inter-language contact. Dialects of a language, by definition, are similar enough to qualify as being a single language by linguistic measures. However, even different languages need not be radically different. The Semitic languages in general show remarkable convergences along different linguistic parameters, long periods of contact certainly being one factor which kept them similar. The example of Aramaic – Arabic contact discussed in 4.2 is but one of many instances of probable convergence via diffusion (see e.g. Owens and Dodsworth 2010). By the same token, when languages in close contact are structurally different, the potential for fairly dramatic diffusional changes are present, the Uzbekistan Arabic case being a clear example, but even in the case of NA idiomaticity borrowing is quite striking in terms of its pervasiveness. In any case, whether there are differences in principle in the social and cognitive factors of bilingual vs. bidialectal borrowing remains an ongoing question.

Abstracting away from all the caveats which are inherent in a comparative discussion of such a wide range of socio-historical linguistic situations, diffusion cannot be seen as a process which ineluctably leads to irregularity. Labov's study elegantly demonstrates that it can indeed do so. His methodological approach is particularly valuable in establishing empirical standards by which change via transmission can be shown to be different from change via diffusion. In the realm of low level phonetics and segmental phonology it could indeed be the case that absolute differences between diffusion and transmission can be established.

On the other hand, in three of the Arabic case studies the effects of diffusion strictly speaking end in greater irregularity only when compared to other Arabic dialects. The dialects, NA, Uzbekistan and Baghdadi Arabic are regular compared to the languages they borrow from. NA idiomaticity, for instance, looks odd in comparison to other Arabic varieties, but not compared to Kanuri, whose system it essentially mimics.

---

18.   Kerswill et al. do suggest that the reduction of diphthongization in central London is simpler than the system it is displacing and that the system is the result of contact with immigrant populations. Hence, one could draw an indirect link between diffusion and simplification. However, incrementation can equally be associated with simplification, as Labov shows in the loss of some lax [a] conditioning environments (e.g. open syllables, certain function words) as the relevant populations moved west in the Midlands area. From this standpoint there is no necessary link between simplification and either transmission or diffusion.

Creole Arabic falls outside the parameters of transmission and diffusion. The sociolinguistic conditions in the second half of the nineteenth century in the southern Sudan were not ones which allowed enough L1 input for a Sudanic or Egyptian Arabic or southern Sudanese languages to become established in the heterogeneous southern camp population. Transmission was impossible, whether from the superstrate, Arabic, or the substrate southern Sudanese languages. If transmission is not an option, diffusion is logically excluded.

If despite the many caveats raised in this final section, one looks for an all-encompassing generalization, the following looks promising. Transmission always results in regularity with change in the system, if it happens at all, incremental. This is demonstrated in variationist terms from the NYC and Northern Cities vowel systems, and also the discourse function of the Arabic verb paradigm. There is no chronological limit on how long the stability may last. Diffusion may, as Labov has it, result in irregularity, though it need not. The 'need not' is sensitive to two very different, hence confounding factors. On the one hand it is sensitive to the linguistic domain it pertains to. Measuring irregularity in some, for instance, semantics or word order, may be inherently impossible, other than via nominal comparison with the basic values one triangulates off of. On the other, from a long term historical linguistic perspective it is generally necessary to rely on outcomes rather than on the ongoing observation of change, where irregularity is more likely to be observable.

I hope to have shown that historically-minded linguists have with Arabic a broad, rich and diverse array of challenges with which to triangulate their studies, and that Arabicists have an interpretive challenge which extends beyond Arabic itself into general questions of the nature of language change.

List of abbreviations

AP	active participle	IND	indicative
C-R	constraint-repair	NA	Nigerian Arabic
DEF	definite	PSSR	possessor
EA	Egyptian Arabic	STA	southern Tunisian Arabic
EmA	Emirati Arabic		

References

Al-Rawi, Rosina. (1990). *Studien zum Arabischen Dialekt von Abu Daby*. Heidelberg: Julius Groos.

Bakker, Peter. (1997). *A Language of our own. The genesis of Michif, the mixed Cree-French language of the Canadian Métis*. Oxford: OUP.

Behnstedt, Peter. (1997). *Sprachatlas von Syrien*. Wiesbaden: Harrassowitz.

Bergsträsser, Gotthelf. (1928 [1977]). *Einführung in die semitischen Sprachen*. Darmstadt: Wissenschaftliche Buchgesellschaft Darmstadt.

Bickerton, Derek. (1981). *Roots of language*. Ann Arbor: Koroma.

Blanc, Haim. (1964). *Communal dialects in Baghdad*. Cambridge Mass.: Harvard Univ. Press.

Blau, Joshua. (2002). *A handbook of Middle Arabic*. Jerusalem: Magnes Press.

Campbell, Lyle. (1993). On proposed universals of grammatical borrowing. In Henk Aertsen and Robert Jeffers (eds.), *Historical Linguistics 1989*, 91–110. Amsterdam: Benjamins. https://doi.org/10.1075/cilt.106.08cam

Campbell, Lyle, Terrence Kaufman and Thomas Smith-Stark. (1986). Meso-America as a linguistic area. *Language* 62: 530–570. https://doi.org/10.1353/lan.1986.0105

Contini, Riccardo. (1999). Le Substrat Araméen en néo-Arabe Libanais: Préliminaires a une enquête systématique. In Marcello Lamberti and Livia Tonelli (eds.), *Afroasiatica Tergestina*, 101–122. Padua: Unipress.

de Landberg, le Comte. (1909). *Etudes sur les Dialectes de l'Arabe Méridionale*. Leiden: Brill.

Erwin, Wallace. (1963). *A short reference grammar of Iraqi Arabic*. Washington D. C.: Georgetown University Press.

Fischer, Wolfdietrich. (1961). Die Sprache der arabischen Sprachninsesl in Uzbekistan. *Der Islam* 36: 232–263. https://doi.org/10.1515/islm.1961.36.3.232

Kaye, Alan. (ed.). (2007). *Morphologies of Asia and Africa*. Winona Lake: Eisenbrauns.

Kerswill, Paul, Eivind Torgersen and Susan Fox. (2008). Reversing drift: Innovation and diffusion in the London diphthong system. *Language Variation and Change* 20: 451–491. https://doi.org/10.1017/S0954394508000148

Kieffer, Charles. (2000). The Arabic speech of Bactria (Afghanistan). In Jonathan Owens (ed.), *Arabic as a minority language*, 181–198. Berlin: de Gruyter. https://doi.org/10.1515/9783110805451.181

Kisseberth, Charles. (2011). Conspiracies. In Marc Osstendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.), *The Blackwell companion to phonology*. Oxford: Blackwells.

Labov, William. (2007). Transmission and diffusion. *Language* 83: 344–387.

Larcher, Pierre. (2001). Moyen arabe et Arabe moyen. *Arabica* 48: 578–609. https://doi.org/10.1163/157005801323163843

Lipiński, Edward. (2000). *The Aramaeans, their ancient history, culture, religion*. Leuven: Peeters.

Macuch, Rudolph. (1982). *Grammatik des samaritanischen Aramäischen*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110862362

Malaika, Nisar. (1959). *Grundzüge der Grammatik des arabischen Dialektes von Bagdad*. Wiesbaden: Steiner.

Malone, Joseph. (1971). Wave theory, rule ordering and Hebrew-Aramaic segolation. *JAOS* 91: 44–66.

Malone, Joseph. (1997). Modern and classical Mandaic phonology. In Alan Kaye (ed.), *Phonologies of Asia and Africa* 1 (pp. 141–159). Winona Lake. Eisenbrauns.

McWhorter, John. (1998). Identifying the Creole prototype: Vindicating a typological class. *Language* 74: 788–818. https://doi.org/10.2307/417003

McWhorter, John. (2002). The rest of the story: Restoring pidginization to creole genesis theory. *Journal of Pidgin and Creole Languages* 17(1): 1–48. https://doi.org/10.1075/jpcl.17.1.02mcw

McWhorter, John. (2007). *Language interrupted: Signs of non-native acquisition in standard language grammars*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195309805.001.0001

McWhorter, John. (2012). Case closed? Testing the feature pool hypothesis. *Journal of Pidgin and Creole Languages* 27: 171–182. https://doi.org/10.1075/jpcl.27.1.07mcw

Mufwene, Salikoko. (2009). Restructuring, hybridization, and complexity in language evolution. In Enoch Aboh and Norval Smith (eds.), *Complex processes in new languages*, 367–401. Amsterdam: John Benjamins. https://doi.org/10.1075/cll.35.23muf

Mufwene, Salikoko. (2010). SLA and the emergence of creoles. *Second Language Acquisition* 32: 359–400. https://doi.org/10.1017/S027226311000001X

Moulton, William. (1954). The stops and spirants of early Germanic. *Language* 30: 1–42. https://doi.org/10.2307/410217

Muraoka, Takamitsu. (1997). *Classical Syriac*. Wiesbaden: Harrassowitz.

Nöldeke, Theodor. (1898). *Kurzgefasste Syrische Grammatik*. Leipzig, Chr. Tauchnitz.

Operstein, Natalie. (2014). Contact-genetic linguistics: toward a contact-based theory of language change. *Language Sciences* 48: 1–15. https://doi.org/10.1016/j.langsci.2014.10.001

Owens, Jonathan. (1994). Nigerian Arabic in comparative perspective. *Sprache und Geschichte in Afrika* 14: 85–175.

Owens, Jonathan. (1996). Idiomatic structure and the theory of genetic relationship. *Diachronica* 13: 283–318. https://doi.org/10.1075/dia.13.2.05owe

Owens, Jonathan. (1997). Arabic-based pidgins and creoles. In Sarah Thomason (ed.), *Contact languages: A wider perspective* (pp. 125–172). Amsterdam: Benjamins. https://doi.org/10.1075/cll.17.08owe

Owens, Jonathan. (1998). *Neighborhood and ancestry: Variation in the spoken Arabic of Maiduguri*. Amsterdam: Benjamins. https://doi.org/10.1075/impact.4

Owens, Jonathan. (2003). Arabic dialect history and historical linguistic mythology. *JAOS* 123: 715–740.

Owens, Jonathan. (2006/2009). *A linguistic history of Arabic*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199290826.001.0001

Owens, Jonathan. (2013a). A house of sound structure, of marvellous form and proportion: an introduction. In J. Owens (ed.), *The Oxford handbook of Arabic linguistics* (pp. 1–22). Oxford: OUP. https://doi.org/10.1093/oxfordhb/9780199764136.001.0001

Owens, Jonathan. (2013b). The historical linguistics of the intrusive **-n* in Arabic and West Semitic. *JAOS* 133: 217–247.

Owens, Jonathan. (2014a). The morphologization of an Arabic creole. *Journal of Pidgin and Creole Linguistics* 29: 232–298. https://doi.org/10.1075/jpcl.29.2.03owe

Owens, Jonathan. (2014b). Many heads are better than one: The spread of motivated opacity via contact. *Linguistics* 52: 125–165. https://doi.org/10.1515/ling-2013-0058

Owens, Jonathan. (2015a). Idioms, polysemy, context: A model based on Nigerian Arabic. *Anthropological Linguistics* 57: 46–98.

Owens, Jonathan. (2015b). Arabic language history and the comparative method. *International Journal of Arabic Linguistics* 1: 1–27.

Owens, Jonathan. (2017). Dia-planar diffusion: Reconstructing early Aramaic-Arabic language contact. In Manuel Sartori, Manuela Giolfo, and Philippe Cassuto (eds.), *Approaches to the history and dialectology of Arabic in honor of Pierre Larcher* (pp. 77–101). Leiden: Brill

Owens, Jonathan and Robin Dodsworth. (2009). Stability in Subject-Verb Word Order: From Contemporary Arabian Peninsular Arabic to Biblical Aramaic. *Anthropological Linguistics* 51: 151–175. https://doi.org/10.1353/anl.0.0014

Owens, Jonathan and Robin Dodsworth. (2017). Semantic mapping: What happens to idioms in discourse. *Linguistics* 55: 641–682.

Owens, Jonathan, Bill Young, Trent Rockwood, David Mehall, Robin Dodsworth. (2010). Explaining Ø and Overt Subjects in Spoken Arabic. In Jonathan Owens and Elgibali eds., *Information structure in spoken Arabic* (pp. 20–60). London: Routledge.

Owens, Jonathan, Robin Dodsworth and Mary Kohn. (2013). Subject expression and discourse embeddedness in Emirati Arabic. *Language Variation and Change* 25: 255–285. https://doi.org/10.1017/S0954394513000173

Paradis, Carole. (1988). On constraint and repair strategies. *The Linguistic Review* 6: 71–97.

Retsö, Jan. (1988). Pronominal suffixes with -n(n)- in Arabic dialects and other Semitic languages. *Zeitschrift für arabische Linguistik* 18: 77–94.

Retsö, Jan. (2000). Kaškaša, t-passives and the dialect geography of ancient Arabia. *Oriente Moderno* 19: 111–118.

Retsö, Jan. (2003). *The Arabs in antiquity*. London: Routledge Curzon.

Ritt-Benmimoun, Veronika, Smaranda Grigore, Jocelyne Owens and Jonathan Owens. (2017). Three idioms, three dialects, one history. *Tunisian and Libyan Arabic Dialect Common Trends – Recent Developments – Diachronic Aspects (Colección Estudios de Dialectología Árabe)* (pp. 43–84). Zaragoza: University of Zaragoza Press.

Rosenthal, Franz. (1961). *A grammar of Biblical Aramaic*. Wiesbaden: Harrassowitz.

Segert, Stanslav. (1997). Old Aramaic Phonology. In Alan Kaye (ed.), *Phonologies of Africa and Asia* (pp. 115–126). Winona Lake, Eisenbrauns.

Segert, Stanislav. (2007). Old Aramaic morphology. In Alan Kaye (ed.), *Morphologies of Asia and Africa* (pp. 121–134). Winona Lake: Eisenbrauns.

Thomason, Sarah and Terrence Kaufman. (1988). *Language contact, creolization and genetic linguistics*. Berkeley: Univ. of California Press.

Versteegh, Kees. (1984). *Pidginization and creolization: The case of Arabic*. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.33

Versteegh, Kees. (1984–86). Word Order in Uzbekistan Arabic and Universal Grammar. *Orientalia Suecana* 33–35: 443–453.

Voigt, Rainer. (2007). Mandaic. In Alan Kaye (ed.), *Morphologies of Asia and Africa* (pp. 149–168). Winona Lake: Eisenbrauns.

Wellens, Ineke. (2005). *The Nubi language of Uganda*. Leiden: Brill.

Windfuhr, Gernot. (2005). Central Asian Arabic: The Irano-Arabic Dynamics of a New Perfect. In Eva Csató, Bo Isaksson and Carina Jahani (eds.), *Linguistic Convergence and Areal Diffusion* (pp. 111–126). London: Routledge.

Vollers, Karl. (1906 [1981]). *Volkssprache und Schriftsprache im alten Arabien*. Amsterdam: Oriental Press.

Zimmermann, Geritt. (2002). *Das Arabische von Buchara zwischen alten Quellen und neuen Forschungergebnissen*. Magisterarbeit: Bayreuth University.

Zimmermann, Geritt. (2009). Uzbekistan Arabic. In Kees Versteegh et al. (eds.), *EALL*, vol. IV, (pp. 612–623). Leiden: Brill.

Online resource

Owens, J. and Jidda Hassan. *In their own voices, in their own words: A corpus of spoken Nigerian Arabic*. http://www.neu.uni-bayreuth.de/de/Uni_Bayreuth/Fakultaeten/4_Sprach_und_ Literaturwissenschaft/islamwissenschaft/arabistik/en/Idiomaticity__lexical_realignment__ and_semantic_change_in_spoken_arabic/Nigerian_Arabic/index.html.

Temporal adverbs of contrast in the Basic Variety of Arabic

Kees Versteegh

University of Nijmegen

This paper applies the model of the Basic Variety developed by Klein & Perdue (1997) and elaborated by Benazzo (2003) to two basic forms of communication in Arabic, Pidgin Madame and Gulf Pidgin Arabic. Benazzo's analysis of the development of temporal adverbs of contrast (resultative *already*; continuative *still*) in the Basic Variety of German, French and English leads to certain predictions about the sequentiality of their acquisition. In the Basic Variety of Arabic the acquisition of these adverbs develops in a different manner. Although their source language does not contain a resultative adverb, both varieties feature such an adverb (*kalas*). This contradicts Benazzo's findings, as does the relatively frequent use of a continuative particle (*bād*) at a very early stage.

Keywords: Basic Variety, Arabic, pidgins, Pidgin Madame, Gulf Pidgin Arabic, adverbs of contrast

1. Basic communication

In his study of language contact in the early colonial Pacific, Drechsel (2014) demonstrates how data from travel accounts and similar sources may be used to reconstruct the linguistic interaction between people speaking different Polynesian languages with strangers. In many cases, they resort to a reduced version of their language in their communication with strangers. In doing so, they adapt their level of speech to the addressees and take into account their presumed proficiency in this particular language type.[1] In their turn, the addressees take the input and process it with their own strategies. This kind of interaction has taken place all over the

---

[1] A well-known instance is that of the Hiri Motu trade language used in Papua New Guinea; speakers of the Austronesian Motu language employed two different reduced versions of their language, one for speakers of related Austronesian languages, and one for speakers of Papuan languages (Dutton 1985).

world and led to the emergence of various kinds of trade and work jargons or early pidgins,[2] such as Gastarbeiterdeutsch (Blackshire-Belay 1991), Namibian Kiche Duits (Deumert 2009), Vietnamese Tây Bôy (Liem 1979), or Indian Butler English (Hosali 2000).

It is a matter of contention what language the new speakers believe they are learning. According to Baker (e.g. 1995; cf. Roberge 2009: pp. 118–119), when two groups without a common language come together, they create a new means of communication that is not based on a pre-existing language as a target for language learning. In my view, this constructive model of pidgin formation confuses product and process. Baker is right in regarding the final product of the encounter as a new variety, but this does not mean that the process itself is one of conscious creation. The learners wish to communicate with the native speakers, but their only information is the initial input they receive. Thus, they have no choice but to take this as their target for learning the new means of communication. Both native speakers and learners are aware of the fact that they are involved in a language learning situation, the former modifying their output in order to facilitate communication, and the latter aiming to reproduce it as faithfully as possible. After prolonged and intensive contact, the learners may come to note the difference between the initial input and the language used by the native speakers among themselves. This may stimulate them to improve their output, but in the process they still target the language of the other group. Not until the learners' variety is recognized as a variety in its own right, can it become a legitimate target for language learning.

Mühlhäusler (1997: p. 138) observes that in highly heterogeneous contexts "in the absence of sufficient overlap and agreement among the speakers of the various jargons in such a situation, universally motivated solutions need to be adopted" (also quoted by Roberge 2009: p. 121). Such views ignore the fact that most linguistic encounters take place between two unequal groups, so that one of them is forced to learn the language of the other group, whether they are traders wishing to sell something to their clients, or domestics on whom a new language is imposed by their employers. It is hard to imagine two groups of equal standing coming together and deciding to communicate by creating an entirely new means

---

2.   This is not the place to enter into a discussion of the various kinds of classification that have been applied to such early forms of communication. In Mühlhäusler's (1997) classification, early stages of communication are called pre-pidgins or jargons. Winford's (2006) earliest stage of pidginization (Stage 1 pidgins), which is characterized by an almost complete absence of functional categories and minimal syntactic structure, is partly identical with Mühlhäusler's pre-pidgins. Sometimes a distinction is made between jargons as individual, and pidgins as collective solutions (Roberge 2009: p. 116).

of communication.[3] Communication results from an attempt to speak a target language that is identifiable as the language of others. This does not mean that the learners are simply copying the input. They immediately start to analyze chunks from the chain of sounds that appear to be useful and relevant to the situation at hand. In this sense, they may be said to 'create' a new code, while at the same time being convinced that they are using their interlocutors' code.

The means of communication resulting from such linguistic encounters is usually called a (trade- or work-related) jargon, and, when it develops into a communal variety, an early pidgin. Although it was not developed to cover cases of emerging trade and work jargons and early pidgins, the Basic Variety model developed by Klein & Perdue (1997) posits an early stage of language acquisition that looks similar to the structure found in them. This model was developed within the framework of the European Science Foundation project *Second language acquisition by adult immigrants*, which focused on learner varieties of German, French, English, and to a lesser degree Dutch and Italian. It dealt with individual learning processes, strictly in terms of their structure. The project's point of departure was that human beings have basic needs that have to be expressed somehow. They need to make clear, for instance, that they want something, which means that they need linguistic means to refer to themselves and to others, since it has to be clear who has to do what for whom. Thus, the central question in Basic Variety research is: how do learners map the available material onto their basic needs.

Klein and Perdue define the Basic Variety as the first variety developed by the new learner of a language, which is systematic in its own right and has its own structure. The lexicon of this Basic Variety contains a number of noun-like and verb-like words, a few adjectives, a number of temporal and spatial adverbs, a negation, a minimal pronominal system, a few quantifiers, and a few prepositions. The syntax operates with a limited number of phrasal and pragmatic constraints, such as the rule that the controller of the action comes first, or the rule that the focus comes last (Klein & Perdue 1997: pp. 312ff.). Explicit marking only takes place "if there is reason to assume that the interlocutor would otherwise reach a false interpretation, and if this false interpretation matters" (Dietrich & Klein 1986: p. 116). Utterances typically consist of uninflected verbs, their arguments, and optional adverbials. There are no free or bound morphemes with purely grammatical functions in the Basic Variety, nor is there any hierarchical system, such as subordination (Klein & Perdue 1997: p. 332).

---

3.   This applies even to cases like Russenorsk, in which the Norwegian and the Russian traders had an equal position (Jahr 1996). One imagines that each group believed they were learning the other's language, hence the mixed nature of the vocabulary.

Later elaborations of the model, for instance by Benazzo (2003) and Benazzo & Starren (2007), introduced a pre-Basic Variety. This earliest stage of L2 learning was also defined as a 'noun-based utterance organization' or as a 'basic lexical stage' (Starren 2006). For the purpose of the present paper, the most important feature identified by the researchers is that at this stage no verbal forms – except frozen ones – are used (Perdue et al. 2002). At the level of the Basic Variety, uninflected verbs begin to occur, and inflected verbs do not appear until the post-Basic Variety.

Winford (2006) claims that the earliest stage of pidgins is more or less identical with what Klein & Perdue call the Basic Variety, and Roberge (2009: p. 109) believes that jargons and Basic Variety should be treated in the same way. According to Roberge they are similar because "[b]oth variety types represent a kind of minimal pragmatic response to communicative exigencies upon initiation of interlingual contact" (2009: p. 116). Yet, the application of the Basic Variety framework to the analysis of these early pidgins is subject to several provisos (see also Owens 2014: pp. 279–284). In the first place, Klein & Perdue themselves (1997: p. 340), while acknowledging the similarities between their Basic Variety, which they see as a form of second language learning, and pidgins, conclude that "all we can say at this point is that there are certainly similarities, but it is quite unclear how far-reaching these are".

A second proviso concerns the status of the 'Basic Variety' itself. The emphasis in Basic Variety research is on features in the learners' speech, rather than on the structure of abstract linguistic systems. This is an improvement compared to the usual approach to the classification of jargons and early pidgins, which seeks to identify discrete varieties. Yet, even the term 'Basic Variety' itself still has essentialist connotations. Klein and Perdue and their colleagues make clear that the Basic Variety represents a stage in language learning rather than a full-fledged variety and that its boundaries are fluid. Nonetheless, it is presented as having certain features that set it apart from other varieties or stages: people are said to go through a pre-Basic Variety, and then to proceed to the Basic Variety, finally reaching the post-Basic Variety. This may be an unintended side-effect of the manner in which the researchers collected their data, by systematically interviewing their informants at certain intervals, each time testing them and assigning them to a specific stage.

In the third place, the model as it was developed by Klein and Perdue is based on a limited set of target languages. It is highly likely that the acquisition of different target languages leads to different characteristics, for instance in the distribution of the parts of speech.[4] We shall see below that in elementary communication

---

4.   Even within the limited set of the ESF project, the researchers point out that there are clear differences in the way target languages are acquired, for instance in the adverbs of contrast, which appear much earlier in learner varieties of French and German than in those of English (see

involving Arabic verbal forms feature conspicuously, while it would be hard to characterize the language proficiency of their users as beyond any pre-Basic stage.

In the fourth place, the data for the ESF project were gathered in specific circumstances, focusing on narratives. The informants were interviewed and requested to retell film fragments. Perdue et al. (2002) state that adults who are learning a new language know already how to assert something, hence they know what to look for in the new language in order to start making assertions. The sociolinguistic setting in which the data for the ESF project were gathered, undoubtedly played a role in this characterization: interviews probably elicit more finite assertions than real life situations do. In early communicative situations, making assertions hardly constitutes the adult learners' most pressing need. Rather, they are primarily interested in expressing their needs and wishes. In situations of a non-narrative nature, even at the very first stages of contact, verbal elements, often in the form of imperatives, probably figure much more frequently than in narrative contexts (Versteegh 2014).

A fifth proviso has to do with the fact that the informants in the ESF project were presented with a 'full' version of the target language: apparently, the interviewers spoke a (simple) variety of the target language and as a rule did not engage in foreigner-directed speech. Such an input differs considerably from the usual practice in early communication and rather resembles a context of formal language teaching. We do not know what happened in the periods between interviews, when the informants were bound to encounter more instances of foreigner-directed, reduced speech. Yet, in the interviews they were treated as learners, who were expected to 'achieve' a next level.

Nonetheless, even with the above provisos in mind and *pace* Klein and Perdue's caution, it is hard to see why the similarity between early L2 acquisition and pidginization should be rejected. Both groups of learners have in common that they intend to speak a target language. In the formal setting of a classroom, early L2 learners are made aware of the deficiency of their speech and do their best to improve its quality. In the type of informal setting in which jargons and pidgins emerge, learners are often discouraged from learning the language and exposed to a reduced version, which they come to believe is the target language. Both groups of learners use similar strategies in communicating, however.

---

below). Moreover, they explain several instances of variation also by referring to the structure of the learners' L1.

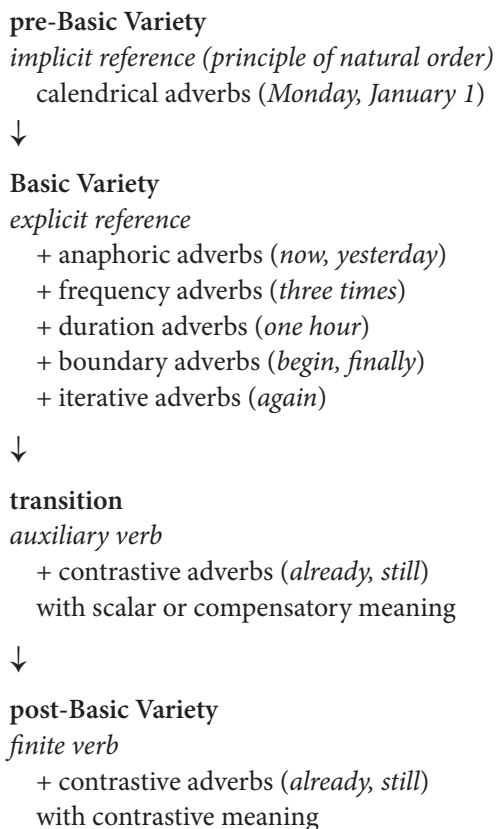## 2.   Expressing temporality in the basic variety

At the pre-Basic Variety stage, almost all utterances consist of nominal and ad-verbial elements, centering around topic/focus messages (Benazzo 2009; see also Benazzo & Starren 2007). Events are positioned on the temporal axis by the speak-ers and understood by the hearers through pragmatic context and general world knowledge. The present time is default and does not need explicitation; for other times calendrical (i.e. deictic) temporal expressions are used. When speaking about the past, the narrative order is that of the natural order of events, which means that the speakers do not need any formal ordering devices.

As long as the learner's variety does not have verbal forms, temporal/aspectual distinctions in the input cannot be expressed morphologically in the learners' out-put. In Klein & Perdue's (1997) model, the first verbal expressions do not make their appearance until the Basic Variety stage; these do not have any inflection (except in a fossilized form), but they do possess an argument structure. Compared to the pre-Basic Variety stage, there are many more temporal expressions, both deictic and anaphorical, and they perform more functions than in the pre-Basic Variety (Benazzo & Starren 2007: p. 138):

–   they situate the event or state on the temporal axis (*yesterday*, *now*)
–   they quantify the frequency and duration of the event (*often, long time*)
–   they introduce temporal limits (*finish*, *first*)
–   they express some aspectual distinctions (habitual, iterative, continuity) (*al-ways*, *again*)
–   they determine the ordering of events (*before*, *after*)

The basic means of communication represented by the Basic Variety seems to work remarkably well without the help of any morphological devices. Nonetheless, at later stages speakers do start to use morphological devices. Benazzo & Starren (2007: p. 138) give two reasons for this development in post-Basic varieties. In the first place, when time reference is lexically-based, there is a higher risk of ambiguity and misunderstandings, because such expressions are easily omitted. In the second place, although some aspectual distinctions can be made without morphological means, other distinctions are bound to remain opaque, such as perfect (topic time after situation time), prospective (topic time before situation time), and progres-sive aspect (topic time in situation time). Benazzo & Starren (2007) argue that verbal morphology is needed because it is more economical in these functions. The exact development of morphological marking is at least partly dependent on the target language, the source language, and probably also on the specific com-bination: Turkish speakers learning Dutch use different strategies than Spanish speakers learning French.

Particles in Basic Varieties, i.e. words that are clearly not noun-like or verb-like, are of several kinds (Benazzo 2003). They are diagnostic in the sense that their acquisition by L2 learners seems to follow a certain sequence (see Figure 1).

pre-Basic Variety
implicit reference (principle of natural order)
   calendrical adverbs (*Monday, January 1*)

↓

Basic Variety
explicit reference
   + anaphoric adverbs (*now, yesterday*)
   + frequency adverbs (*three times*)
   + duration adverbs (*one hour*)
   + boundary adverbs (*begin, finally*)
   + iterative adverbs (*again*)

↓

transition
auxiliary verb
   + contrastive adverbs (*already, still*)
   with scalar or compensatory meaning

↓

post-Basic Variety
finite verb
   + contrastive adverbs (*already, still*)
   with contrastive meaning

Figure 1. Stages in the acquisition of temporal adverbs (after Benazzo 2003)

According to Benazzo, temporal adverbs of contrast such as *already*, *still*, *yet* are acquired at a rather late stage, not because they represent cognitively complex notions – which they do – but because they are co-dependent on verbal morphology, which typically emerges only at the post-Basic Variety stage. In her account, at the pre-Basic Variety stage, temporal and aspectual relations are indicated by pragmatic principles such as the Principle of Natural Order, which stipulates that the default interpretation of a narrative is that events are reported in the order in which they occurred, and by calendrical adverbs (*May 1*; *1985*). The only particles that are used at this stage are additive (*also*) or restrictive (*only*, *except*) ones. At the next stage, that of the Basic Variety, lexical means are used to refer to tense and aspect. At this stage, new temporal adverbs appear that indicate boundaries, position, frequency, or duration. At this stage, narrated events are still more or less simultaneous to the time specified by lexical means. In addition to the additive and restrictive adverbs used in the pre-Basic Variety, iterative adverbs, such as *again*, appear.

Productive morphology does not appear until the post-Basic Variety stage. It is only at this stage that resultative (*already*) or continuative (*still*) adverbs of contrast may be used with verbal predicates. Benazzo reports that in the Basic Variety, even before the appearance of finite verbs, learners of French and German start using *déjà/schon* with adverbs of frequency and position, as in (1).

(1) *schon   drei   mal   mit   ihm*
     already three time with 3M.SG
     '[I was] already three times with him'   (Benazzo 2003: p. 199)

But in such cases, *déjà/schon* is used with a scalar meaning ('It was not the case that I had seen him only once, but I saw him three times'), having scope over the adverb. With a resultative meaning *déjà/schon* appear at a later stage. For learners of English, *already* does not appear before the production of finite verbs and only occurs in the speech of advanced learners.

During the transition between Basic Variety and post-Basic Variety German *noch*, French *encore* and English *still* make their first appearance, as in (2).

(2) *this man still to take some apples*
     'This man is taking some apples'   (Benazzo 2003: p. 203)

In this intermediate stage, when there are already verbal forms, but no finite predicates, these adverbs of contrast may be used as compensatory devices contributing to the temporal or aspectual interpretation of the utterance (Benazzo 2003: pp. 204–205). In the absence of unambiguous verbal morphology, *already/déjà/schon* indicate past tense or perfect aspect, while *still/encore/noch* refer to imperfective aspect. It is impossible to use them with a purely temporal meaning before the appearance of verbal morphology, when they can have scope over the predicate. The use of the negative contrastive adverb (*not yet*) is achieved only by the most advanced learners (Benazzo 2003: pp. 194–195).

When full-blown verbal morphology becomes available, the importance and frequency of temporal adverbs decreases since some of their functions are taken over by finite verbs. On the other hand, the possibilities of inserting the adverbial expression increase. In the pre-Basic Variety, the only possible constructions for additive or restrictive particles, for instance, are Particle – X or X – Particle (where X indicates any other constituent), but at later stages, when infinite or finite verbs come to be used, other positions become available (Perdue et al., 2002).

Both Benazzo (2003) and Perdue et al. (2002) emphasize the fact that the use of particles strongly depends on the stage acquired by the learner. In the pre-Basic Variety, additive and restrictive particles are possible because these operate on entities (NPs), which are typically available at this stage. In the Basic Variety, iterative particles may occur because they quantify over bounded events, which can be

signalled independently of verbal morphology. For *already* and *still*, however, which compare similar events in different time intervals, verbal morphology is needed in order to make the comparison possible. Starren (2006) shows that the marking of prospective and perfect aspect is a necessary condition for the functioning of these two adverbs of contrast.

3.   Temporal adverbs in Basic Varieties of Arabic

Corresponding to the Basic Varieties of German, French, and English L2 speakers, there are some varieties of Arabic that have emerged in similar circumstances. Gulf Pidgin Arabic (also known as Urdu Pidgin Arabic) arose as the principal means of communication between Arab employers and foreign workers in Saudi Arabia and the Gulf states. Many of the speakers of this variety come from South Asia and have an Indian language as their L1 (Naess 2008; Al-Moaily 2008; Bakir 2010). Pidgin Madame is the name given by Bizri (2010) to a pidginized variety of Arabic that is spoken in Lebanon between Sri Lankan domestics having Sinhalese as their mother tongue and their Lebanese employers.[5] These varieties exhibit an extreme degree of variability, but at the same time they serve as a conventionalized means of communication, which the users regard as learnable.

The variety of Arabic these learners are exposed to is extremely reduced. Bizri (2010: pp. 147–154) illustrates this by citing instances of native speakers 'translating' their own utterances into a reduced version when addressing a Sri Lankan girl. In fact, the availability of a full version of the language to be learnt can hardly be regarded as self-evident. Owens (2014: p. 283) believes that learners are always confronted with normal discourse, which means that they have at their disposal a full set of inflectional material, which they start to analyze immediately. The question is, however, on what basis the learners construct this system. Bizri's data show quite convincingly that there is a huge difference between the native speakers' normal discourse and the way they address the learners. Thus, the learners do not have enough data to develop inflectional paradigms.

The Basic Varieties of Arabic do not have any grammatical markers for the expression of tense or aspect, but they do have verbal forms with an argument structure, as in (3a):

---

5.   In quoting examples from Bizri's material I have retained as much as possible the transcription of the original; note that *ṃ* denotes a velar nasal and *ṭ* a retroflex dental. In glossing the examples from Pidgin Madame and other varieties I have followed the intended rather than the etymological meaning of the forms used; thus, a sentence like *ana yešūp* will be glossed as '1SG see', rather than '1SG see.IMPERF.3M.SG'.

(3) a. *badēṃ ana srabi neskape*
         then   1SG drink Nescafe
         'Then, I drink a cup of coffee'   (Bizri 2010: p. 176)
     b. *awwal ana ma yešūp huwa*
         first   1SG NEG see   3M.SG
         'I had never seen him before'   (Naess 2008: p. 87)

The verbal forms used in these examples are uninflected forms without agreement or tense/aspect markers. Etymologically, they derive from a variety of verbal forms in the lexifier. Gulf Pidgin Arabic forms often derive from imperatives in the target language, a peninsular Arabic dialect. In Pidgin Madame, many of the verbal forms go back to Lebanese Arabic feminine imperatives (Versteegh 2014), for instance the form *srabi* in (3). Other forms are found as well, often second person singular, but also other verbal forms, such as the third person singular *yešūp* in (3).[6]

The utterances in Gulf Pidgin Arabic and Pidgin Madame are temporally anchored, either on the basis of contextual indications, or by pragmatic principles, such as the Principle of Natural Order. The discourse context usually suffices to allow the speakers to distinguish between events having happened earlier on, events happening right now, and events expected or wished or intended to occur in the future. If the natural sequence of events is not clear, chronology may be indicated with the help of temporal contextualizers, such as *badēṃ* 'then' (< Arabic *baʿdēn* 'afterwards'). The default time is the present, while the past may be referred to with *awwäl* (< Arabic *ʾawwal* 'first') or *abel* (< Arabic *ʾabl* 'before'); if the reference to the present needs to be highlighted, *hälla* (< Arabic *hallaʾ* 'now') is used. There is no grammaticalized way to refer to the future. Most references to the future in Bizri's texts deal with the expression of desire or intention, or with obligations. Verbs meaning 'to want' may signal the first appearance of grammaticalized marking, as in (4).

(4) *hälla sway baddik nhottu bank*
     now   little want put   bank
     'Now, I'll put away some [money] in the bank'   (Bizri 2010: p. 237)

Since diachronic materials of the kind used in the ESF project are rarely available for Arabic pidgins, it is impossible to set up a sequence for the acquisition of particles in these varieties in the way Benazzo did for German, French, and English L2 varieties. It is, however, possible to observe the aggregated use of adverbs in a corpus

---

6. Owens (2014: p. 289, n. 59) allows for the use of some imperatives in the Sudanese army camps, where Juba Arabic was employed, but claims that this possibility is otherwise unsubstantiated. Yet, it does not take much imagination to believe that in this context, the predominant use of imperatives was in fact a stark reality.

in order to determine which developmental stage it represents. Table 1 lists all adverbs that occur in the texts provided by Bizri (2010), showing that Pidgin Madame possesses the full array of adverbs that is commonly assigned to the pre-Basic and Basic Variety. Adverbs without a temporal connotation include an intensifier, a negation, and some deictic particles. Temporal adverbs include calendrical adverbs (not shown in Table 1), anaphoric adverbs, and adverbs of frequency and duration. On the whole, therefore, the system in Pidgin Madame corresponds to what may be expected on the basis of Benazzo's and others' inventory for the Basic Variety.

Table 1. Number of occurrences of adverbs in Pidgin Madame (based on the corpus of texts I–XII in Bizri 2010)

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	
TEMPORAL													
awwäl 'at first'	1					2	3	1	3		1	1	12
abel 'before'		4			4	7							15
hälla 'now'	7	10		1	4		3	2	1		3	5	36
lyōm 'today'	1												1
SEQUENCE													
badēṃ 'afterwards'	9	13	5	2	11		11				2	10	63
ADDITIVE													
kamēṃ 'also'	7	2		3	3				1		2		18
RESTRICTIVE													
bäs 'only'	2				5		1		2		3	1	14
FREQUENCY													
kilyōm 'everyday'		2	3		1	1	2				3	1	13
marra 'time, turn'					1	1	1					1	4
CONTRASTIVE													
bād 'still, yet'	6	7			2				4		4		23
NEGATIVE													
ma 'not'	5	16			14	2	5	3		1	12	3	61
la 'not'							3	2	2	5			12
no 'not [Engl.]'	2	5			5		1		2		3	2	20
mat 'not'											6		6
mes 'not'		1			1							2	4
LOCAL													
hōn 'here'	1	11			3		2		1	1	2	4	26
honīke 'there'							2					5	7
MANNER													
hēk 'so'	1	7	5		19		2	1	7	1	11	1	55
dogre 'quick'					4		1					1	6
ṭīr 'much, very'		2	1		3		1				2	1	10
DURATION													
sämēn 'long time'												1	1

The list in Table 1 shows that Pidgin Madame exhibits one major exception to Benazzo's (2003) observations about the acquisition of adverbs in the (pre-)Basic Variety. As we have seen above, Benazzo concludes that adverbs meaning *already* or *still, yet* are not acquired until the last stage of L2 acquisition. This conclusion does not seem to hold for the acquisition of these particles in Pidgin Madame.

   With respect to the resultative adverb of contrast, represented in English by the particle *already*, there are indications that at least in some of the Arabic Basic Varieties the particle *kalas* has this function. The etymological origin of *kalas* is the Arabic adverb *ḫalāṣ* 'that's it!, done!', from the verb *ḫalaṣ* 'to be finished', which in contemporary dialects expresses completion. In Cairene Arabic, for instance, *ḫalāṣ* may be used both pre- and postverbally (Woidich 2006: p. 273).[7] Postverbally, it indicates completion or finality, as in (5); preverbally, there may be a connotation of 'definitely, really', as in (6).

(5) a. *gahhizu*   *ḫalāṣ*
         make.ready.PFV.3PL finish
         'Are they finished with preparing the trousseau?' (Woidich 2006: p. 273)

     b. *d-ana*   *mutti*   *ḫalāṣ*
         here-1SG   die.PFV.1SG finish
         'I am really exhausted'   (Woidich 2006: p. 273)

(6) a. *ʾana ḫalāṣ ʾakalt*
         1SG   finish eat.PFV.1SG
         'I've had enough'   (Badawi & Hinds 1986: p. 260)

     b. *hiyya ḫalāṣ baʾit*   *kiwayyisa*
         3F.SG   finish become.PFV.3F.SG good
         'She has become a good girl now'   (Badawi & Hinds 1986: p. 260)

The expletive use of *ḫalāṣ* in native speech, often accompanied by a hand gesture indicating finality or completion of a job, is at the basis of its function in foreigner-directed speech as a completion marker.

   In Gulf Arabic, the particle *kalaas* is used postverbally, where it may be analyzed as a completive marker, as in (7). But it also occurs in preverbal position, where it functions as a resultative adverb of contrast, as in (8).

(7) *ʾatbuk kalaas laham šilli*
     cook   COMPL meat   take.out
     'When the cooking is done, I take out the meat'   (Bakir 2010: p. 212)

---

7.   Note that Bruweleit (2015: pp. 102ff.) does not mention any verb or particle with the sense of 'already' in Beirut Arabic; anteriority in this dialect is indicated by the use of the perfect verb with an auxiliary *kān*.

(8) *’inta kalaas waddi fuluus*
     2SG   already give   money
     'Have you sent the money?'   (Bakir 2010: p. 213)

In the few occurrences of *kalas* in Pidgin Madame, a similar use is observed, as in (9) and (10).

(9) *oktik itnēn kalas   jawās*
     sister two   already marriage/marry
     'The two sisters were already married'   (Bizri 2010: p. 126)

(10) *hälla kullu kalas   sēwe bil bēt*
     now all   already do   in   house
     'Everything has already been done in the house'   (Bizri 2010: p. 127)

In (9), *jawās* could be interpreted as a nominal form, in which case *kalas* would convey a verbal meaning ('to finish marriage', i.e. 'to marry'). But in (10), the only possible interpretation seems to be that it is used as an adverb of contrast, equivalent to English *already*. The form *kalas* does not occur postverbally in the data about Pidgin Madame provided by Bizri.

   The relatively infrequent use of a resultative adverb in Arabic Basic Varieties may perhaps be explained by the structure of the target language, because the input for the pidgin learners does not contain any clearly recognizable resultative particles. In Classical Arabic *qad* is used as a resultative particle, but in the vast majority of contemporary Arabic dialects there is no comparable word. Perfect aspect is often indicated with different means, for instance, by the use of the participle, or by periphrastic expressions with a verb meaning 'to finish'. The acquisition of *kalas* at an early stage in Arabic Basic Varieties differs from the sequence set up by Benazzo for Basic Varieties of German, French, and English, where it appears only at a later stage.

   Concerning the continuative adverb (equivalent with English *still*, *yet*), too, the situation in Pidgin Madame is strikingly different from that in German, French, and English as an early L2. In Classical Arabic, *ba‘du* is used as a continuative particle, in particular in combination with a negation in the sense of 'not yet'. In some contemporary dialects, for instance in Beirut Arabic, it is frequently used for what Bruweleit (2015: p. 99) calls 'still-situations', as in (11).[8]

---

8. In Egyptian Arabic, a similar function is carried out by the particle *lissa*, which according to Woidich (2006: pp. 283–289) means 'still' with stative and inchoative verbs; in combination with negative *ma-* it means 'not yet' (2006: p. 349).

(11) *kãnit   ba'da   'am-tibki   b-ūḍt-a   lamma*
      be.PFV.3F.SG still   CONT-weep.IPFV.3F.SG in-room-3F.SG when
      fãt   ba-yya
      come.by.PFV.3M.SG   father-3F.SG
      'She was still weeping in her room when her father entered'
                                                              (Bruweleit 2015: p. 99)

In Bizri's corpus there are twenty-two utterances contain the particle *bād* (or *bad*).
The meaning of *bād* in these utterances is not identical. We have seen above that
Benazzo (2003) distinguishes between a scalar (additive) and a contrastive meaning
for English *still*/French *encore*, the scalar meaning being acquired earlier than the
contrastive one. In the majority of utterances in Pidgin Madame, the meaning of
bād seems indeed to be scalar, as in (12)–(14).

(12) *badēṃ ana bad nēn säher   sirlanka*
      then   1SG BAD two month Sri Lanka
      'Then, I stayed another two months in Sri Lanka'   (Bizri 2010: p. 234)

(13) *bad wehde badde jībi*
      BAD one   want bring
      '[He] wants to bring another one'   (Bizri 2010: p. 183)

(14) *enti bād jäws*
      2SG BAD marry/husband
      'You have to marry again'   (Bizri 2010: p. 158)

There are, however, five cases in which the meaning is clearly contrastive, in four of
which the particle is used with a negation, corresponding to English *not yet*/French
pas encore, as in (15) and (16); the one positive example is in (17).

(15) *bād no   äkel*
      BAD NEG   eat
      '[I] have not yet eaten'   (Bizri 2010: p. 161)

(16) *bād no   estegel kalas*
      BAD NEG [Eng] work   finish
      '[I] have not yet finished working'   (Bizri 2010: p. 161)

(17) *pi   hayda bado sərire*
      EXS this   BAD   young
      'She is still young'   (Bizri 2010: p. 169)

Studies of Gulf Pidgin Arabic cannot be compared directly to Bizri's study of Pidgin
Madame since they do not contain complete texts, but only example sentences. In
Naess (2008: p. 41, 57, 74, 90) *bād* occurs four times; in all cases it has additive/
scalar meaning, as in (18).

(18) *mafi ijlis yistokol bād*
      NEG rest work   BAD
      'We can't rest! We must work more'   (Naess 2008: p. 74)

At the very least, this shows that the particle occurs in this related variety, albeit only in the simpler function of an additive or scalar adverb of contrast. In view of the nature of the material, nothing much can be said about its frequency.

4. Conclusion

The comparison between the material in the ESF project, which is remarkably homogeneous, and the material in Bizri's corpus underscores the fact that different language pairs are needed to confirm or reject any hypothesis concerning the universality of developments in the Basic Variety. In the literature about the Basic Variety, differences in the output of learners of German, French and English are explained either as the effect of the different structure of the target languages, or as the effect of interference from the source language, the learners' L1. The case of Arabic (pre-)Basic Varieties shows that the study of other language pairs may indeed lead to different observations about the sequentiality of acquisition of these adverbs.

The effect of the structure of the target language may even go deeper than the use of temporal adverbs of contrast and involve the distribution of parts of speech. According to all descriptions of pre-Basic and Basic Variety, the first stage is exclusively noun-based, whereas verbs do not make their first entry until the Basic stage. Verbal forms that do appear before this stage must be considered to be frozen material and are probably best analyzed as nominal forms. Yet, in the contact between speakers of Arabic and those of Sinhalese or Urdu, which led to the emergence of Pidgin Madame and Gulf Pidgin Arabic, things turn out to be different. In referring to both actions and states, the speakers of Pidgin Madame and Gulf Pidgin Arabic quite often use verbal forms. The phonetic shape of these verbal forms varies considerably, but they do not exhibit any agreement markers.[9] Yet, although some of the

---

9. In this respect, Pidgin Madame and Gulf Pidgin Arabic resemble the Australian pidgin described by Dench (1998). This Pidgin Ngarluma was used in intertribal contact in the North West Cape, and also functioned as a jargon in contacts with foreigners. In 1876, two rescued castaways wrote an account of their contacts with the Aborigines, with whom they stayed for a period of no more than six months, during which they learned Pidgin Ngarluma at an elementary level. Apparently, they used (uninflected) verbal forms, predominantly for basic notions such as coming, eating, drinking etc.

speakers seem to have acquired a high degree of fluency in handling this variety,[10] its structure is commensurate with the level of a pre-Basic variety.

The difference between the Arabic (pre-)Basic Varieties and the results from the ESF project with respect to the acquisition of contrastive adverbials may partly be explained by a difference in research design between the two fields. Researchers of the Basic Variety held long sessions with their informants, in which the latter were exposed to the full variety of the language, whereas this was not the case in Pidgin Madame, nor in Gulf Pidgin Arabic. As we know from Bizri's account, the Sri Lankan domestics in Lebanon have practically no other contact than with their female employers, who expose them to a rudimentary form of Arabic only. Both the Sri Lankan domestics and the Southasian migrant workers have to adapt quickly to this type of verbal communication with their employers, otherwise they are out of a job. In both cases, imperatives are the main device to convey the wishes of the employer. For the learners, these forms are the point of departure for their own speech. The data show that in spite of the absence of inflected verbal forms Pidgin Madame and Gulf Pidgin Arabic have a full array of adverbials, including the resultative and the continuative particle that in other linguistic encounters appear associated with a post-Basic Variety.

List of abbreviations

COMPL	completive		NEG	negation
EXS	existential		PFV	perfective
F	feminine		PL	plural
IPFV	imperfective		SG	singular
M	masculine			

---

10. Recently, the first movie featuring a Gulf Pidgin Arabic speaker was produced by Ray Had-dad. It was entitled *Being Sayed Rasoul: A day in the life of a Pakistani labor worker,* and por-trayed a Pakistani who had worked for many years as a taxi driver in the Gulf. This documentary was a huge success when it was shown at the Abu Dhabi film festival in 2012; see https://vimeo.com/56636702. The protagonist of this moving story has developed his own variety of Arabic, which he speaks with the utmost confidence. Likewise, some of the Sri Lankan girls in the doc-umentaries *Maid in Lebanon* I and II shot by Carol Mansour in 2005 and 2008, even though their proficiency is highly limited, speak their variety of Pidgin Madame without hesitation (see http://h-net.msu.edu/cgi-bin/logbrowse.pl?trx=vx&list=h-gender-mideast&month=0810&week=b&msg=Y7pNy9/EvZjiqsbx3mnPWw&user=&pw=). The conversation of the speakers of Juba Arabic in Manfredi's texts in the online corpus of African varieties of Arabic <http://corpa-froas.tge-adonis.fr/Archives/ListeFichiersELAN.php> likewise provide interesting examples of fluent pidgin spoken by people who are obviously used to speaking this variety. See Clements (2003), who distinguishes between halting speech and fluent speech production.

Bibliographical references

Al-Moaily, M. (2008). *A data-based description of Urdu Pidgin Arabic*. M.A. thesis, University of Newcastle.

Badawi, E. & Hinds, M. (1986). *A dictionary of Egyptian Arabic*. Beirut: Librairie du Liban.

Baker, P. (1995). Motivation in creole genesis. In P. Baker (ed.), *From contact to creole and beyond* (pp. 3–15). London: University of Westminster Press.

Bakir, M. J. (2010). Notes on the verbal system of Gulf Pidgin Arabic. *Journal of Pidgin and Creole Languages* 25, 201–228. https://doi.org/10.1075/jpcl.25.2.01bak

Becker, A. & Veenstra, T. (2003). The survival of inflectional morphology in French-related creoles. *Second Language Acquisition* 25, 283–306.

Benazzo, S. (2003). The interaction between the development of verb morphology and the acquisition of temporal adverbs of contrast: A longitudinal study in French, English and German L2. In C. Dimroth & M. Starren (eds.), *Information structure, linguistic structure and the dynamics of language acquisition* (pp. 187–210). Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/sibil.26.09ben

Benazzo, S. (2009). The emergence of temporality: From restricted linguistic systems to early human language. In R. Botha & H. de Swart (eds.), *Language evolution: The view from restricted linguistic systems* (pp. 21–57). Utrecht: LOT.

Benazzo, S. and Starren, M. (2007). L'émergence de moyens grammaticaux pour exprimer les relations temporelles en L2. *Acquisition et interaction en langue étrangère* 25, 129–157.

Bizri, F. (2010). *Pidgin Madame: Une grammaire de la servitude*. Paris: Geuthner.

Blackshire-Belay, C. (1991). *Language contact: Verb morphology in German of foreign workers*. Tübingen: G. Narr.

Bruweleit, S. (2015). *Aspect, tense and action in the Arabic dialect of Beirut*. Leiden: E.J. Brill. https://doi.org/10.1163/9789004287549

Clements, J. C. (2003). The tense-aspect systems in pidgins and naturalistically learned L2. *Studies in Second Language Acquisition* 25, 245–281. https://doi.org/10.1017/S0272263103000111

Dench, A. (1998). Pidgin Ngarluma: An indigenous contact language in North Western Australia. *Journal of Pidgin and Creole Languages* 13, 1–61. https://doi.org/10.1075/jpcl.13.1.02den

Deumert, A. (2009). Namibian *Kiche Duits*: The making (and decline) of a Neo-African language. *Journal of Germanic Linguistics* 21, 349–417. https://doi.org/10.1017/S1470542709990122

Dietrich, R. and Klein, W. (1986). Simple language. *Interdisciplinary Science Review* 1:2, 110–117.

Drechsel, E. J. (2014). *Language contact in the early colonial Pacific: Maritime Polynesian Pidgin before Pidgin English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139057561

Dutton, T. (1985). *Police Motu: Iena sivarai (Its story)*. Port Moresby: The University of Papua New Guinea Press.

Hosali, P. (2000). *Butler English: Form and function*. Delhi: B.R. Publishing Corporation.

Jahr, E. H. (1996). On the pidgin status of Russenorsk. In E. H. Jahr & I. Broch (eds.), *Language contact in the Arctic: Northern pidgins and contact languages* (pp. 107–122). Berlin & New York: Mouton de Gruyter.

Klein, W. & Perdue, C. (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13, 301–347. https://doi.org/10.1191/026765897666879396

Liem, N. D. (1979). Cases and verbs in Pidgin French in Vietnam (Tay Boi). In P. Mühlhäusler et al. (eds.), *Papers in Pidgin and Creole Linguistics* 2 (pp. 217–246). Canberra: Australian National University.

Mühlhäusler, P. (1997). *Pidgin and creole linguistics*. 2nd ed. London: University of Westminster Press.

Naess, U. G. (2008). *'Gulf Pidgin Arabic': Individual strategies or a structured variety? A study of some features of the linguistic behaviour of Asian migrants in the Gulf countries*. M.A. thesis, University of Oslo.

Owens, J. (2014). The morphologization of an Arabic creole. *Journal of Pidgin and Creole Languages* 29, 232–298. https://doi.org/10.1075/jpcl.29.2.03owe

Perdue, C., Benazzo, S. & Giuliano, P. (2002). When finiteness gets marked: The relation between morpho-syntactic development and use of scopal items in adult language acquisition. *Linguistics* 40, 849–890. https://doi.org/10.1515/ling.2002.032

Roberge, P. T. (2009). The creation of pidgins as a possible window on language evolution. In R. Botha & H. de Swart (eds.), *Language evolution: The view from restricted linguistic systems* (pp. 101–137). Utrecht: LOT.

Starren, M. (2006). Temporal adverbials and early tense and aspect markers in the acquisition of Dutch. In V. van Geenhoven (ed.), *Semantics of acquisition* (pp. 219–244). Berlin: Springer. https://doi.org/10.1007/1-4020-4485-2_9

Versteegh, K. (2014). Pidgin verbs: Imperatives or infinitives?. In I. Buchstaller, A. Holmberg & M. Almoaily (eds.), *Pidgins and creoles beyond Africa-Europe encounters* (pp. 141–169). Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/cll.47.07ver

Winford, D. (2006). Reduced syntax in (prototypical) pidgins. In L. Progovac, K. Paesani, E. Casielles, & E. Barton (eds.), *The syntax of non-sententials* (pp. 283–307). Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/la.93.13win

Woidich, M. (2006). *Das Kairenisch-Arabische: Eine Grammatik*. Wiesbaden: Otto Harrassowitz.

On the relationship between Arabic Foreigner Talk and Pidgin Arabic

Andrei Avram

University of Bucharest

The paper compares morphosyntactic and lexical features of the Arabic Foreigner Talk register to those of four Arabic-lexifier pidgins, Pidgin Madame, Jordanian Pidgin Arabic, Romanian Pidgin Arabic, and Gulf Pidgin Arabic. The comparative overview identifies a relatively significant number of features which Arabic Foreigner Talk shares with all or with at least some of these Arabic-lexifier pidgins. The paper proposes an account in terms of a feedback relationship whereby Arabic Foreigner Talk and Pidgin Arabic reinforce one another in the occurrence of these features.

Keywords: Arabic Foreigner Talk, Pidgin Arabic, morphology, syntax, vocabulary

1. Introduction

The aim of the present paper is to assess the potential role of the Arabic Foreigner Talk register in the emergence of pidginized varieties of Arabic. The empirical data analyzed are illustrative of the morphosyntactic and lexical features of Arabic Foreigner Talk (henceforth FT) and of four Arabic-lexifier pidgins: Pidgin Madame (henceforth PM), Jordanian Pidgin Arabic (JPA), Romanian Pidgin Arabic (RPA), and Gulf Pidgin Arabic (GPA).

As noted by Al-Sharkawi (2007: p. 117), "research in registers modified by native speakers represents a new field of inquiry in Arabic linguistics". The studies published so far are representative of only a few varieties of Arabic FT: Egyptian Arabic FT (Al-Sharkawi 2005, 2007, 2010), Jordanian Arabic FT (Tweissi 1990), Kuwaiti Arabic FT (Dashti 2013), Lebanese Arabic FT (Haraty & al. 2007; Bizri 2010), and Saudi Arabic FT (Al-Ageel 2015). The Arabic FT samples discussed in this paper are drawn from studies employing a variety of methods for data collection. These include: recordings of spontaneous speech (Tweissi 1990; Al-Sharkawi 2005, 2007;

Haraty & al. 2007; Bizri 2010; Dashti 2013); controlled elicitation (Tweissi 1990; Al-Sharkawi 2005, 2007; Al-Ageel 2015); volunteer reports (Al-Sharkawi 2005, 2007, 2010); self-reports (Al-Sharkawi 2005, 2007, 2010); reports in the media (Naylor 2008, Zacharias 2010, Al Hameli 2013). Representations of Arabic FT in e.g. TV series and films have not been included since these do not reflect actual usage (for a discussion, see Al-Sharkawi 2007: p. 118–120).

The data on PM are from Bizri (2005), Haraty & al. (2007), Bizri (2005, 2009, 2010). The sources for JPA data are Al-Salman (2013) and Al-Haq & Al-Salman (2013). The data on RPA are from Avram (2010, own corpus). For GPA two types of data are included: from general descriptions of GPA (Smart 1990; Wiswall 2002; Avram 2014, 2016) as well as from studies on GPA as spoken in several countries: Saudi Arabia (henceforth SA) – Almoaily (2008), Al-Azraqi (2010), Albakrawi (2012), Almoaily (2013), Alghamdi (2014), Almoaily (2014), and Al-Zubeiry (2014); Kuwait (K) – Salem (2013); Qatar (Q) – Bakir (2010); Oman (O) – Næss (2008), Alshuaimi (2011); additional data are from my own corpus. The data are authentic, since examples, from e.g. the media, cartoons, TV series, films, etc., illustrative of how native speakers of Arabic imagine GPA to be or of their attempts at imitating it, have not been included.

The examples from both Arabic FT and the varieties of Pidgin Arabic considered appear in a uniform system of transliteration.[1] For GPA the country in which a particular feature has been documented is also specified.

The paper is organized as follows. Section 2 focuses on morphosyntactic features attested both in Arabic FT and in the four Arabic-lexifier pidgins. Section 3 is concerned with lexical features of Arabic FT and of the varieties of Pidgin Arabic considered. In Section 4, I outline a possible account of the similarities observed between Arabic FT and these Arabic-lexifier pidgins. The findings are summarized in Section 5.

2. Morphosyntax

2.1 'Two' + singular noun

According to Al-Sharkawi (2007: p. 119), an "avoided element from the recorded native-speaker [Egyptian Arabic] FT data is the dual ending *-ēn*". As shown below, the dual is replaced by "*'itnēn* 'two' followed by the noun", which "appears sometimes in the singular and sometimes in the plural" (Al-Sharkawi 2010: p. 234):

---

1. The symbol <þ> stands for a PM consonant "to be placed between the bilabial stop /p/ and the labio-dental fricative /p/" (Bizri 2010: p. 15).

(1) a. *ʔitnēn kitāb*   (Al-Sharkawi 2010: p. 234)
         two   book
         'two books'
     b. *ʔitnēn laʕʕība*   (Al-Sharkawi 2007: p. 119)
         two   dogs
         'two dogs'

The use of 'two' followed by a noun in the singular is also attested in Kuwaiti Arabic FT (Dashti 2013: 78):

(2) *ṭalʔi aṯnēn diǧāǧa*   (Dashti 2013: p. 79)
     get out two   chicken
     'prepare two chickens'

The absence of the dual marker and its replacement by a noun phrase consisting of the numeral 'two' and a noun in the singular is a common feature of all the Arabic-lexifier pidgins considered, as illustrated by the following examples from RPA (3), JPA (4), RPA (5), and GPA (6):

(3) *nēn yōm*   (Bizri 2010: p. 116)
     two day
     'two days'

(4) *wi-tnēn ukti*   (Al-Salman 2013: p. 41)
     and two sister
     'I have one brother and two sisters.'

(5) *itnen dinar*   (Avram 2010: p. 23)
     two   dinar
     'two dinars'

(6) a. *tanēn marah*   SA (Alghamdi 2014: p. 14)
         two   time
         'twice'
     b. *itnen paučča*   K (Salem 2013: p. 108)
         two   child
         'two children'
     c. *tinēn ʔusbū*   Q (Avram 2014: p. 17)
         two   week
         'two weeks'
     d. *isnēn sana*   O (Avram 2014: p. 17)
         two   year
         'two years'

2.2   Plural marker 'all'

Arabic FT is also characterized by the absence of the Arabic means of marking plurality (internal plural or plural suffixes). Dashti (2013: p. 77), for instance, states that Kuwaiti Arabic FT uses "the word /killə/, meaning 'all of it' to indicate the plural".

(7)   *ǧībi*   *hāḏa* *ǧanṭə* **killə**   *māl* *āna*   (Dashti 2013: p. 77)
       bring.IMP.2F.SG DEM bag   all of it POSS 1SG
       'bring my bags'

A similar observation is made by Bizri (2010: 116) with respect to Pidgin Madame: "in the absence of a morphological plural marker […] *kello* "all", placed before a noun in the singular, assumes this function".

(8)   **kello** *bēbi*   (Bizri 2010: p. 116)
       all   baby
       'babies'

The use of 'all' to mark plurality is documented for RPA as well:

(9)   *sayara* **kulu-kulu**   (Avram 2010: p. 23)
       car   all all
       'cars'

Note that when used as a plural marker in RPA, the quantifier *kulu* always appears in its reduplicated form *kulu-kulu* and it is placed in postnominal position. GPA appears to be another Arabic-lexifier pidgin making the same use of 'all'. Although no examples are provided, Albakrawi (2012: p. 129) writes that GPA as recorded in Saudi Arabia may also mark plurality by means of the quantifier *kullu* 'all'.

2.3   Omission of the definite article

The omission of the definite article is reported with reference to Kuwaiti Arabic FT (Dashti 2013: p. 73).

(10)   *Tati, eš   fīč   dāḫil* Ø *seyāra.*   (Dashti 2013: p. 73)
       Tati   what in 2F.SG inside   car
       'Tati, what's wrong with you, it's in the car.'

Although not explicitly mentioned, Saudi Arabic FT also appears to exhibit this feature:

(11) *kīs Ø ǧīb*   (Al-Ageel 2015: p. 117)
      bag   bring
      'Bring the bag'

On currently available evidence, the omission of the definite article is attested in three of the Arabic-lexifier pidgins considered. Consider the following examples from PM (12), RPA (13), and GPA (14) respectively:

(12) *ana no   teftahi Ø bēb*   (Bizri 2010: p. 99)
      1SG NEG open   door
      'I don't open the door'

(13) *Ø inǧiner   šuf inte*   (Avram 2010: p. 24)
          engineer see 2SG
      'The engineer sees you.'

(14) a. *Ø kafīl   fī sawwi ǧinǧāl*   Q (Bakir 2010: p. 217)
              sponsor FI make   quarrel
          'The sponsor quarrels [with me]'
      b. *Ø Muškil   eš?*   O (Alshuaimi 2011)
              problem what
          'What's the problem?'

2.4 Masculine singular form of adjectives

Self reports discussed in Al-Sharkawi (2010: p. 235) mention the lack of gender of gender agreement, with the masculine singular used as the default form:

(15) *ʔintī **kuwayyis** kitīr*   (Al-Sharkawi 2010: p. 235)
      2F.SG good.M.SG much
      'you are really good'

Similarly, only the masculine singular form of adjectives is used in Kuwaiti Arabic FT (Dashti 2013: p. 80), as in (16), which illustrates the lack of both gender and number agreement:

(16) *šīli hāḏa   ǧiwāti **qadīm***   (Dashti 2013: p. 80)
      take DEM.M.SG shoes   old.M.SG
      'keep all these old shoes'

Al-Ageel (2015: p. 128) also mentions the lack of gender agreement in Saudi Arabic FT. In (17), the masculine form of the adjective is used even though the noun it modifies is feminine:

(17) **asfar**   ǧubna   waḥid lī   ǧībī   (Al-Ageel 2015: p. 128)
      yellow.M cheese.F one   for-1SG bring
      'Bring me one yellow cheese [sandwich]'

Three of the Arabic-lexifier pidgins at issue attest to the use of the masculine singular adjective as the default form. These are JPA (18), RPA (19), and GPA (20):

(18) *hunak fulūs   muš **kuwayes***   (Al-Salman 2013: p. 68)
      there   money NEG good
      'The salaries there are not good.'

(19) *Hada sayara **zen?***   (Avram, own corpus)
      DEM   car   good
      'Is this car good?'

(20) *mumkin hiya   **tābān***   O (Næss 2008: p. 41)
      maybe   3F.SG tired
      'Maybe she's tired.'

2.5   Exclusive use of independent pronouns

The samples of Kuwaiti Arabic FT (21), Saudi Arabic FT (22) and Lebanese Arabic FT (23) illustrate the use of independent pronouns instead of pronominal suffixes:

(21) *hāḏa mū   māl   **āna***   (Dashti 2013: p. 75)
      DEM NEG   POSS 1SG
      'These are not mine.'

(22) *ana maktab*   (Al-Ageel 2015: p. 126)
      1SG office
      'my office'

(23) a. *beybi **ente** kbīre*   (Bizri 2010: p. 217)
          baby 2SG big
          'your child will grow up'
      b. ***ente** hazbend ēmtan   mūte?*   (Bizri 2010: p. 149)
          2SG husband when   die
          'When did your husband die?'

Note that when it encodes the possessor, the independent pronoun may occur in prenominal position, as in (22) and (23b), even though this does not reflect the order of constituents in Arabic.

The use of independent pronouns to the exclusion of pronominal suffixes is amply documented for all the pidginized varieties of Arabic under discussion, as shown in the examples below from PM (24), JPA (25), RPA (26), and GPA (27):

(24) ***ana bēt***   (Bizri 2010: p. 117)
      1sg house
      'my house'

(25) *ʔana bisāid **huwwa***   (Al-Salman 2013: p. 66)
      1sg help 3sg
      'I help him'

(26) *Inǧiner šuf **inte***   (Avram 2010: p. 24)
      engineer see 2sg
      'The engineer sees you.'

(27) *māmā yabi ʔanā*   Q (Avram 2014: p. 17)
      Madam want 1sg
      'Madam wants me.'

2.6   Masculine singular form of demonstratives

The neutralization of gender and number distinctions in demonstratives is attested in Kuwaiti Arabic FT (Dashti 2013: p. 83), in which only the masculine singular form is used. In addition to Kuwaiti Arabic FT (28), this is also true of Saudi Arabic FT (29):

(28) ***hāḏa***   *malābis   killə   niḏīf*   (Dashti 2013: p. 80)
      DEM.M.SG clothes.PL all of it clen
      'These clothes are clean.'

(29) *Ana **haḏa**   ġurfa   147*   (Al-Ageel 2015: p. 133)
      1sg DEM.M.SG room.F.SG 147
      'I am in room 147'

The consistent use of the masculine singular of the demonstrative, as an invariant form, is widely recorded in both RPA (30) and GPA (31):

(30) *giv **hada** sikina la ani*   (Avram 2010: p. 30)
      give DEM knife to 1sg
      'give me that knife.'

(31) a.   ***haza** nafarat zen.*   K (Salem 2013: p. 108)
           DEM men   good
           'These men are good.'
      b.   *asān   **hāda** mama   kalām arabi*   O (Avram 2014: p. 29)
           because DEM madam speak Arabic
           'Because the madam [only] spoke Arabic'

2.7   Invariant form of verbs

In Arabic FT the form of the verb is subject to reduction and simplification. Al-Sharkawi (2007: p. 119) quotes volunteer reports according to which in Egyptian Arabic FT "native speakers of Arabic delete the imperfect 2nd and 3rd person prefixes on the stem of the verb", as in the following example:

(32)   *ʔinta*   ***yi-šrab***   (Al-Sharkawi 2007: p. 119)
       2M.SG   3M.SG-drink
       'you drink'

Another possibility, mentioned in self-reports, is "the use of the masculine second and third person imperfective […] for both genders" (Al-Sharkawi 2010: p. 235):

(33)   *ʔintī*   ***ti-ktib***   (Al-Sharkawi 2010: p. 235)
       2F.SG   2M.SG-write.IPFV
       'you write'

With respect to Lebanese Arabic FT, Al-Sharkawi (2007: p. 119) reports the use of the feminine imperative as a finite verb "when addressing the housemaids". More recently, Bizri (2010: p. 148) writes that "*Madame* opts without hesitating for the imperative forms":[2]

(34)   *ana* ***rūḥe***   *māma*   (Bizri 2010: p. 148)
       1SG go.IMP.F.SG mother
       'I went to my mother's'

According to Al-Ageel (2015: p. 117), in Saudi Arabic FT "two forms of a verb can be used alternatively within the same conversation by the same speaker". In fact, different forms of the same verb – the imperative and the imperfective – may also be found in one and the same sentence:

(35)   *kīs* ***jīb***   *baʕdēn* *zabadī* ***yi-ǧīb***   *entī*
       bag bring.IMP.2M.SG then   yogurt 3M.SG-bring.IPFV 2SG
       'Bring [me] the bag and then the yogurt'   (Al-Ageel 2015: p. 117)

As for Kuwaiti Arabic FT, Dashti (2013: p. 71) notes a preference for the use of the imperfective form of verbs:

(36)   *Soma āna* ***yi-sawwi***   *talifūn*   *ams*   (Dashti 2013: p. 72)
       Soma 1SG 3M.SG-make.IPFV   telephone yesterday
       'Soma, I telephoned yesterday'

---

2.   *Madame* is the term designating the female employer of foreign domestic workers.

Conversely, in the Arabic FT used in the United Arab Emirates (Al Hameli 2013), the perfective may occur as the invariant form of the verb:

(37) *ana* **fakkr** (Al Hameli 2013)
1SG think.PFV.3M.SG
'I thought'

Summing up, Arabic FT is characterized by a strong tendency towards using what is essentially an invariant form of the verb, regardless of aspect, person, gender, and number. The same holds for Arabic-lexifier pidgins. Although the origin of the form of the verb is a matter of some debate in the literature (for discussion, see e.g. Bizri 2012; Versteegh 2014a), there is consensus that verbs typically occur in an invariant form, although this is subject to variation, depending on the particular Arabic-lexifier pidgin. Thus, according to Bizri (2010: p. 74), PM makes use of verbs etymologically derived from feminine forms, most of which in the imperative.[3] In JPA forms derived from the imperfective appear to predominate (Al-Salman 2013; Al-Haq & Al-Salman 2013), whereas RPA displays a preference for those derived from imperative forms (Avram 2010). Finally, both the imperfective and the imperative have yielded forms used in GPA (Bakir 2010: pp. 206–209; Avram 2014: p. 18).

2.8 Light verb 'make' + noun/adjective/verb

Arabic FT exhibits structures consisting of the light verb 'make' + noun/adjective/verb. The example below is from Egyptian Arabic FT:

(38) **ʕamalt** **sūra** **šaxsiyya** (Al-Sharkawi 2010: p. 234)
make.PFV-1SG picture personal
'I had a picture taken of myself.'

The *Madames* in Bizri (2010) also make use of such structures in their Lebanese Arabic FT:

(39) *ana bellēl* **ʕamele talifon** (Bizri 2010: p. 148)
1SG in the night make telephone
'I will phone her tonight'

Wiswall's (2002) and Dashti's (2013) examples illustrate the occurrence of these structures in Kuwaiti Arabic FT as well:

---

3. Bizri's informants were all housemaids, interacting mostly with their female employers, which accounts for the occurrence of feminine forms of verbs.

(40)　a.　*āna* **yi-sawwi**　　**talifūn**　　*ams*　　　　　　　(Dashti 2013: p. 72)
　　　　　1SG　3M.SG-make　telephone　yesterday
　　　　　'I phoned yesterday'
　　　b.　*anta* **sawwi**　**fakkar**　　　　　　　　　　　　　　(Wiswall 2002)
　　　　　2SG　make　　think
　　　　　'you think'

The extensive use of 'make' + noun/adjective/verb structures is documented for PM, JPA, and GPA, as shown in (41), (42), and (43) respectively:

(41)　*ana* **sewi**　**akel**　　　　　　　　　　　　　　　　　(Bizri 2009: p. 9)
　　　　1SG　make　food
　　　　'[if] I cook'

(42)　*bādu* **sawwi**　**zadīd** *hada*　*mudīr*　　　　　　　(Al-Salman 2013: p. 42)
　　　　then　make　　new　　DEM　manager
　　　　'then the manager renews it'

(43)　a.　*lēš*　*māmā*　*māfī*　**sawwi**　**tabdīl**　　　　　Q (Bakir 2010: p. 219)
　　　　　why　Madam　NEG FI　make　　change
　　　　　'why doesn't Madam change [it]'
　　　b.　*ana* **sawwi**　**nadīp**　　　　　　　　　　　　　　O (Næss 2008: p. 91)
　　　　　1SG　make　　clean
　　　　　'I clean'
　　　c.　*bukra*　　*ana* **sawe**　**tasel**　　*anta*　　　SA (Avram, own corpus)
　　　　　tomorrow　1SG　make　contact　2SG
　　　　　'I will contact you tomorrow'

2.9　Reliance on context or on time adverbials

The examples in Dashti (2013) and in Al-Ageel (2015) show that both Kuwaiti Arabic and Saudi Arabic FT rely either on the context or on time adverbials for tense and aspect marking. Consider the following example from Kuwaiti Arabic FT:

(44)　*anā*　*yi-gūlič*　　　　　*misāʕ*　　　　　　　　　　(Dashti 2013: p. 71)
　　　　1SG　3M.SG-tell-2F.SG　from hour
　　　　'I told you an hour ago'

All the four Arabic-lexifier pidgins under discussion share this characteristic, as seen in the examples below, from PM (45), JPA (46), RPA (47), and GPA (48):

(45)　**abel**　*ma*　*þi*　*si.*　　　　　　　　　　　　　　　(Bizri 2010: p. 122)
　　　　before　NEG　FI　thing
　　　　'There was nothing before.'

(46) *abu   ʔana mūt* **gabul** *tamantās sana*   (Al-Salman 2013: p. 62)
father 1SG   die   before eighteen year
'My father died eighteen years ago.'

(47) *Leš rua* **dilwati?**   (Avram 2010: p. 26)
why go   now
'Why are you leaving now?'

(48) a.   *tadrīb* **awwal** *šwayy*   SA (Albakrawi 2012: p. 129)
practice first   a little
'I practiced a little before.'

   b.   **ʔamis**   *ʔanā yabi …*   Q (Bakir 2010: p. 206)
yesterday 1SG   want
'Yesterday, I wanted …'

   c.   *Shuwaya, jay,   ana fi kalām*   O (Alshuaimi 2011)
a little   come 1SG FI speak
'After a while, I came and talked.'

2.10   Multifunctional *fi/fī*

The samples of Saudi Arabic FT in Al-Ageel (2015) illustrate the occurrence of *fī* as a predicative copula:

(49) **kwayes** *mafī   baʕdēn*   (Al-Ageel 2015: p. 122)
good   NEG FĪ then
'It will not be good, then'

This use of *fī* is also reported by Brockett (1985: 24), as a feature of what he calls "debased and slang usage" in the Khabūra dialect of Omani Arabic:

(50) *mā* **fī** **zayn** *hēde*   (Brockett 1985: p. 24)
NEG FĪ good DEM
'this isn't good'

The same use of *fī/fi* is attested in both JPA (Al-Salman 2013: p. 68) and GPA (Bakir 2010: pp. 216–217; Avram 2012: pp. 40, 43, 45, 47–48, 50, and 52; Avram 2013; Al-Shurafa 2014: p. 18; Bakir 2014: pp. 420–42). Consider the following examples from JPA (51) and GPA (52):

(51) *kullu fī   gāli*   (Al-Salman 2013: p. 68)
every FĪ expensive
'everything is expensive'

(52)   a.   *fi ahsan*   SA (Avram 2012: p. 20)
            FI good
            'it's alright'
       b.   *ʔinta fī maǧnūn*   Q (Bakir 2010: p. 216)
            2SG FĪ crazy
            'Are you crazy?'
       c.   *Bēlad fi zein.*   O (Alshuaimi 2011)
            country FI good
            'Hometown is good.'

Kuwaiti Arabic FT also exhibits *fi* + verb structures:

(53)   *anta fī fakkar*   (Wiswall 2002)
       2SG FĪ think
       'you think'

Such structures also occur in the Khabūra variety of Omani Arabic studied by
Brockett (1985):

(54)   *baʕdayn fī šill fir-rās*   (Brockett 1985: p. 24)
       then   FĪ take in head
       'then he takes it to the head'

The co-occurrence of *fī/fi* and verbs is attested in JPA (55) and GPA (56):

(55)   *bēbi fi nōm hassa*   (Al-Salman 2013: p. 65)
       baby FI sleep now
       'The baby is sleeping now.'

(56)   a.   *ana fi gul inta tāl bet*   SA (Avram 2014: p. 23)
            1SG FI say 2SG come house
            'I told you to come [to my] place.'
       b.   *ana fi maʔlum*   K (Salem 2013: p. 109)
            1SG FI know
            'I know.'
       c.   *ʔinta fī yaskit*   Q (Bakir 2010: p. 217)
            2SG FĪ be silent
            'You keep quiet.'
       d.   *ana fi sugul hamstašar sana*   O (Alshuaimi 2011)
            1SG FI work fifteen year
            'I've been working for fifteen years.'

Note that the grammatical status of *fī/fi* as used in such cases is controversial in
the literature on GPA. The controversy essentially focuses on whether *fi/fii* is a
particle (Al-Azraqi 2010: p. 169), a predication marker (Bakir 2010: pp. 215–219;

Al-Shurafa 2014: p. 19; Bakir 2014: pp. 433–434), a verbal predicate marker (Avram 2012: pp. 54–55, 2013) or a copula (Potsdam & Alanazi 2014: p. 28).

2.11 Omission of prepositions

The samples of Lebanese Arabic FT in Bizri (2010) demonstrate that both the *Madames* and their friends frequently omit prepositions, in particular those marking direction or location:

(57) *bokra   bēt   kello rūḥe Ø Bayrūt*   (Bizri 2010: p. 154)
      tomorrow house all   go   Beirut
      'Tomorrow, we'll all go to Beirut'

The frequent omission of directional or locative prepositions appears to be typical of Saudi Arabic FT as well:

(58) *ana Ø haḏa ġurfa 147*   (Al-Ageel 2015: p. 133)
      1SG   DEM room 147
      'I am in room 147'

With respect to PM, Bizri (2010: 130) specifies that it is "characterized by the absence of morphemes marking the directive, the ablative, the locative".

(59) a. *Ø sawdiya   rāhet*   (Bizri 2010: p. 130)
          Saudi Arabia go
          'I went to Saudi Arabia.'
     b. *kullu Ø sirlanka   ǧip*   (Bizri 2010: p. 130)
          all   Sri Lanka bring
          'I had brought everything from Sri Lanka.'

As seen in the following examples, the omission of directional and locative prepositions is also characteristic of JPA (60) and of GPA (61):

(60) *bādēn […] izi Ø Zōrdan*   (Al-Salman 2013: p. 63)
      then   go   Jordan
      'then came to Jordan'

(61) a. *Ø Bangladesh bas bangla*   SA (Almoaily 2013: p. 186)
          Bangladesh only Bengali
          'in Bangladesh [there is] only Bengali'
     b. *ana Ø Waziristan*   UAE (Avram, own corpus)
          1SG   Waziristan
          'I am from Waziristan'

c.	*bādēn*	*sīr*	Ø	*dikān*	*araf*	O (Næss 2008: p. 58)
	then	go		shop	know	

'[she] also knows how to go to the shop'

d.	*ʔanaa*	*māfi*	*rūḥ*	Ø	*sīnema*	Q (Bakir 2010: p. 207)
	1SG	NEG FI	go		cinema	

'I don't go to the cinema.'

3.   Vocabulary

3.1   Lexical polysemy

According to Al-Ageel (2015: p. 127), Saudi Arabic FT is characterized by "the lack of unified vocabulary system". This appears to be true of other varieties of this register and accounts for, among others, the fact that one of the basic strategies identified by Al-Sharkawi (2007: p. 117) is "the explanation of lexical items". In addition, the vocabulary of Arabic FT is limited in size. Bizri (2010: p. 151), for instance, writes that "*Madame* confines herself to the limits of the lexical stock of the Sri Lankan [maid]".

   A consequence of the reduced size of the vocabulary is lexical polysemy, whereby lexical items may undergo semantic extension. A case in point is discussed by Bizri (2010: p. 151), who states that "*Madame* has understood that in order to say "send something somewhere" it would be better to say […] "something goes somewhere"", as illustrated below:

(62)   *ente maṣāre   kello rūḥe serlanka*   (Bizri 2010: p. 151)
       2SG money   all   go   Sri Lanka
       'Did you send all your money to Sri Lanka?'

Instances of lexical polysemy obtaining via semantic extension are found in the available samples of PM and GPA. In the following example from PM, the verb 'to give' is used with the meaning of 'to save':

(63)   *ana **atet**   masare*   (Bizri 2010: p. 178)
       1SG give   money
       'I saved money'

Similarly, in GPA, the verb 'to sit' may also mean 'to rest', 'to stay', and 'to live':

(64)   a.   ***yiğlis*** *andel sandūg māl   cash*   O (Avram 2014: p. 30)
            sit   at   box   POSS cash
            'I was sitting at the cash register'

    b.   *mafi*   **iğlis**   O (Avram 2014: p. 30)
          NEG FI rest
          '[We] can't rest'

    c.   **yiğlis** *bēt*   O (Avram 2014: p. 30)
          sit   house
          'stay at home'

    d.   *ana* **iğlis** *hina*   O (Avram 2014: p. 30)
          1SG sit   here
          'I live here.'

3.2   Circumlocutions

The limited available vocabulary also accounts for the fact that users of Arabic FT resort to paraphrases or to circumlocutions. With respect to PM, for instance, Bizri (2010: p. 150) observes that "*Madame* very frequently resorts to paraphrases".

(65) **ḥoṭṭe**   *bāba* **nār** *baʕdēn*   (Bizri 2010: p. 150)
      put.IMP.2F.SG father fire   then
      'then [they] put father on fire' [= 'incinerate']

The same strategy is employed in PM (66) and GPA (67):

(66) *wehde byestegel þi siyara, baddik   tsēwe siyara, baddik   tsēwe hēk*
      one   work   in car   you want make car   you want make thus
      'one works on cars, [he] makes cars, [he] does things like this' [= mechanic]
                                                                 (Bizri 2010: p. 102)

(67) a.   *omur kabīr*   SA (Almoaily 2013: p. 174)
         age   big
         'elderly'

      b.   *āti   halīb*   O (Avram 2014: p. 32)
         give milk
         'breastfeeds'

As seen in (66), circumlocutions can be at times rather lengthy and convoluted renderings of a particular meaning.

3.3   English lexical items

As noted by Tweissi (1990: p. 308) and Al-Sharkawi (2007: p. 120), the vocabulary of Arabic FT includes foreign lexical items, in particular of English origin. In both Egyptian and Jordanian Arabic FT these are used to explain the meaning

of Arabic words unknown to non-native speakers. In Lebanese Arabic FT, these English lexical items are found either instead of or as synonyms of Arabic words (see Bizri 2010: p. 152), i.e. they appear to function as nonce borrowings, presumed to be better understood than their Arabic equivalent. The following examples are illustrative of Lebanese Arabic FT as used by an interviewer of domestic workers (68a) and by a *Madame* (68b):

(68) a. *inti **what** ištigil*   (Haraty & al. 2007)
          2SG what 3M.SG.work
          'What do you do?'

      b. ***oḫtik***   *bi lebnēn? ente **sister** hälläʔ lebnēn?*
          sister-2F.SG in Lebanon 2SG sister now   Lebanon
          'Is your sister now in Lebanon?'   (Bizri 2010: p. 152)

Al-Ageel (2015: p. 127) also notes the occurrence of English words in Saudi Arabic FT. Consider the example below:

(69) ***chicken** waḥid abġa*   (Al-Ageel 2015: p. 128)
      chicken one   want
      'I want one chicken'

The use of English lexical items is widely attested in all the four Arabic-lexifier pidgins investigated. According to Bizri (2005: p. 63), in PM the English word "appears either together with its Arabic equivalent (to facilitate the possibilities of comprehension", as in (70a), or "instead of its equivalent", as in (70b):

(70) a. *rah wēn   ana **ämbäsi** saḅāra*   (Bizri 2010: p. 98)
          go where 1SG embassy embassy
          'Where did I go? [Well,] to the embassy.'

      b. *barrad **clean**, himmam   **clean**, killo.*   (Haraty & al. 2007)
          fridge clean bathroom clean all
          'I clean the house, clean he fridge, the bathroom, all of it'

In JPA, English words are frequently used instead of Arabic-derived lexical items:

(71) *hada momgin **finiš***   (Al-Salman 2013: p. 42)
      DEM maybe   finish
      'this might end'

The following pair of examples from RPA shows that the same speaker may use alternately Arabic or English words as synonyms:

(72) a. *Inte* *šogol* *zen.*   (Avram 2010: p. 23)
         2SG work good
         'You work well.'

      a′. *Inte* *no* **work** *maku* *fulus*   (Avram 2010: p. 31)
         2SG NEG work NEG be money
         'If you don't work, [you] have no money.'

In GPA, Arabic- and English-derived lexical items may occasionally co-occur in the same sentence:

(73) **tanēn second** *čiko*   O (Avram 2014: p. 32)
      two   second child
      '[my] second child'

Far more frequent is the alternate use of Arabic and English words as synonyms, which may occur in the speech of the same user of GPA:

(74) a. *yalla*   **bas** *ḫalas*   SA (Avram 2014: p. 32)
         alright but finish
         'alright, but that's it'

      b. *ana* *malūm* **but** *šuf*   SA (Avram 2014: p. 32)
         1SG know   but see
         'I know, but look'

4. Discussion

Several characteristics of the morpho-syntax (see Section 2) and vocabulary (see Section 3) of Arabic FT match those found in the varieties of Pidgin Arabic considered. These are set out in Table 1 (for GPA, also specified are the countries in which the occurrence of a feature has been documented).

   One issue which needs to be addressed is how can the existence of these features both in Arabic FT and in the four Arabic-lexifer pidgins examined be accounted for. A rather widespread opinion among native speakers of Arabic, in particular in the Gulf region, is that pidginized Arabic emerges because non-Arabs are essentially exposed to the Arabic FT register. Naylor (2008), for instance, quotes a professor of linguistics at the United Arab Emirates University as saying that "we [= Arabs] think that people from other languages will not understand us if we speak normally". According to Al Hameli (2013), in the United Arab Emirates "it is common to hear Arabs speaking in broken Arabic [to non-Arabic speakers]". Al-Bargi (2014) reports on the views expressed by both native speakers of Arabic and by foreign nationals in Saudi Arabia: a businessman believes that "it is much easier to use

Table 1. Shared features in Arabic FT and Arabic-lexifiers pidgins

Feature	Arabic-lexifier Pidgin (Country)
'two' + singular noun	PM, JPA, RPA, GPA (SA, K, Q, O)
plural marker 'all'	PM, RPA, GPA (SA)
omission of definite article	PM, JPA, RPA, GPA (SA, K, Q, O)
masculine singular form of adjectives	JPA, RPA, GPA (O)
exclusive use of independent pronouns	PM, RPA, GPA (SA, Q, O)
masculine singular form of demonstratives	RPA, GPA (K, O)
invariant form of verbs	PM, RPA, GPA (SA, K, Q, O)
light verb 'make' + noun/adjective/verb	PM, JPA, GPA (SA, Q, O)
reliance on context or time adverbials	PM, JPA, RPA, GPA (SA, Q, O)
predicative copula *fi*	JPA, GPA (SA, Q, O)
verbal predicate marker *fi*	JPA, GPA (SA, K, UAE, Q, O)
omission of prepositions	PM, JPA, GPA (SA, K, UAE, Q, O)
lexical polysemy	PM, GPA (O)
circumlocutions	PM, GPA (SA, O)
English lexical items	PM, JPA, RPA, GPA (SA, K, Q, O)

broken Arabic grammar, pronunciation and vocabulary […] with non-Arabs than to speak […] in plain Arabic" (Al-Bargi 2014); a Filipino nurse states that "many patients speak to me in a very telegraphic language with distorted pronunciation and vocabulary"; a foreign waiter (nationality not disclosed) complains that "customers] will […] speak with a non-Saudi accent".

Such metalinguistic comments are somewhat reminiscent of the so-called "baby-talk" theory and of the FT hypothesis. In a well-known and oft-quoted passage, Bloomfield (1933: p. 472) writes that "speakers of a lower language may make so little progress in learning the dominant speech, that the masters, in communicating with them resort to 'baby-talk'", which is "the master's imitation of the subjects' incorrect speech". Bloomfield (1933: p. 472) goes on to state that "the subject, in turn, deprived of the correct model, can do no better than to acquire the simplified 'baby-talk' version of the upper language". As for the potential role of FT in the genesis of pidgins, Ferguson (1971: p. 121) thinks that it "may serve as an incipient pidgin". On this view, "the initial source of the grammatical structure of a pidgin is the more or less systematic simplification of the lexical source language […] in the foreigner talk register" (Ferguson 1971: p. 121). Moreover, Ferguson (1971: p. 121) sees FT as a more significant factor in the emergence of pidgins "than the grammatical structure of the language(s) of the other users of the pidgin" (see also Ferguson & DeBose 1977).

Assessing the potential role of Arabic FT in the formation of Pidgin Arabic is no easy task. As mentioned in Section 1, only a few varieties of Arabic FT have

been documented and there are very few studies based on recordings of spontaneous speech. Nonetheless, currently available evidence does permit a tentative, though provisional, assessment. Firstly, like other varieties of FT (see Mühlhäusler 1997: pp. 97–98; Sebba 1997: p. 90), Arabic FT appears to be variable in many respects. Note, however, that the distribution of the features considered in Sections 2 and 3 cuts across varieties of Arabic FT. Secondly, a feature may be more robustly attested in a particular variety of Arabic FT, which may impact on its frequency of occurrence in the local Pidgin Arabic. Thirdly, it is certainly not the case that Arabic FT is the source of all the morphosyntactic and lexical features attested in Pidgin Arabic. Finally, the phonology of Pidgin Arabic cannot be traced back to Arabic FT. Al-Sharkawi (2007: p. 120), for instance, concludes that "in the natural data, no articulatory modifications were recorded" and that "native speakers [of Arabic] did not resort to any alteration of the phonological features of sounds".[4] A similar point is made by Bizri (2010: p. 151), who writes that "on the phonological level, the Arabic of *Madame* is very resistant". Also, there is no evidence of such phonological adjustments in the samples of Kuwaiti Arabic FT (Dashti 2013) and of Saudi Arabic FT (Al-Ageel 2015). Forms used in the United Arab Emirates Arabic FT, such as *fakkr* 'think' and *khabr* 'say', show that consonant gemination and the velar voiceless fricative /x/ are preserved in Arabic FT, even though these do not occur in GPA.

The suggestion is advanced here that there is a feedback relationship holding between Arabic FT and the Arabic-lexifier pidgins considered. In other words, Arabic FT and these varieties of pidginized Arabic reinforce one another in the occurrence of certain features. According to Mühlhäusler (1997: p. 102), "the importance of foreigner talk in Pidgin formation appears to be restricted to relatively early stages of development", i.e. before the stabilization stage (in the classification proposed by Mühlhäusler 1997: p. 6). This would accord with the fact that, on most accounts (Bizri 2010 on PM; Al-Haq & Al-Salman 2013, Al-Salman 2013 on JPA; Avram 2010 on RPA; Næss 2008, Avram 2013, 2014 and 2016, Alghamdi 2014 on GPA; Versteegh 2014b), the four varieties of Pidgin Arabic considered in this paper are in the jargon/pre-pidgin stage (in terms of the classification proposed by Mühlhäusler (1997: p. 6). On the other hand, pidgins can also contribute to the formation of the FT register. In their discussion of the relationship between FT and the Atlantic creoles, den Besten & al. (1995: p. 97) write that FT is "often modeled on pidgins, so that the latter may erroneously be thought to have emerged out of the former". A feedback relationship between Arabic FT and Pidgin Arabic cannot therefore be ruled out.

---

4. Al-Sharkawi (2007: p. 118) mentions the fact that such changes, e.g. "the regular shift of /ḥ/ into /x/" or the substitution of "/ʿ/with a glottal stop/ʾ/", are only found in "Foreigner Talk in movies".

As shown in Sections 2 and 3, the Arabic FT register may have contributed to the occurrence in Pidgin Arabic of the morphosyntactic and lexical features discussed in Sections 2 and 3. The question arises, however, whether Pidgin Arabic also influences Arabic FT. Although not necessarily conclusive, there appears to be some evidence in this respect. For instance, light verb 'make' + verb structures (Sections 2 and see 2.8) occur more frequently in Kuwaiti Arabic FT than in GPA (Wiswall 2002), presumably under the influence of the latter. Wiswall (2002) also notes an overuse of *fī* + verb (see 2.10) in Kuwaiti Arabic FT compared to GPA. What Brockett (1985: p. 25) calls "superfluous *fī*", i.e. *fī* used as predicative copula and as verbal predicate marker, may also be a characteristic of Omani Arabic FT. According to Brockett (1985: p. 25), "*fī* is used in profusion by Indians and Pakistanis when speaking Arabic" and "perhaps this is imitated by Omanis when talking to them and to other foreigners, thinking that by doing so they will be more easily understood". Bizri (2010: p. 150) notes that in Lebanese Arabic FT, *Madame* resorts to circumlocutions (see 3.2) "using the same procedure as the Sri Lankan [housemaid]". It has also been suggested that Pidgin Arabic may even influence the locally spoken dialect of Arabic. With respect to GPA, Almoaily (2013: p. 184) writes that "there are potential pieces of evidence for lexical as well as morphological effects of GPA on GA [= Gulf Arabic]". On the morphosyntactic level, Almoaily (2013: p. 184) comments on examples such as *ṯalaṯa riyal* 'three riyal' that "the reason why the GPA-like pluralization is used in GA with the word *riyal* […] is probably because GA speakers mostly have daily monetary transactions […] with GPA speakers". Almoaily (2013: p. 184) also puts forth the hypothesis that lexical items such as *sīdah* 'straight' and *nafar* 'person' may have been borrowed by Gulf Arabic not directly from Urdu, but rather via GPA.

5.   Conclusion

This paper has examined a number of morphosyntactic and lexical features found both in Arabic FT and in four Arabic-lexifier pidgins, PM, JPA, RPA, and GPA.

It is certainly not claimed that all features of these varieties of Pidgin Arabic are also found in the Arabic FT register. In addition to the phonology, which clearly reflects the influence of the first languages of the users of the Arabic-lexifier pidgins discussed (on RPA, see Avram 2010: pp. 21–22; on GPA, see Avram 2014: p. 3, Avram 2017: pp. 132–133), there are also morphosyntactic features which seem not to occur in Arabic FT, but which can be traced back to substratal influence (on GPA, see Bakir 2010: p. 221), and others which may be the outcome of incipient grammaticalization (on GPA, see e.g. Avram 2012: pp. 54–55, 2014: pp. 35–36, 2017: p. 142). A possible "conspiracy of factors" needs also to be taken into account:

features found in Pidgin Arabic may be the result of convergent influences of different sources (on GPA, Avram 2014: pp. 36–37, 2017: pp. 143–147).

The suggestion that Arabic FT played a role in the emergence of Arabic-lexifier pidgins is not new in itself. Tosco & Manfredi (2013: p. 510), for instance, write that "certainly the influence of foreigner talk was important in the genesis of GPA", but do not make reference to any specific features. Avram (2014: pp. 34–36) discusses several morphosyntactic features of GPA that might conceivably be attributed to the influence of the Arabic FT register. However, this paper is the first systematic overview of morphosyntactic and lexical features shared by Arabic FT and the four Arabic-lexifier pidgins at issue. Also, rather than assuming that Arabic FT is the source of Pidgin Arabic or, conversely, that the former is modeled on the latter, it is suggested that a feedback relationship holds between the two, which influence one another.

List of abbreviations

DEM	demonstrative		NEG	negative
F	feminine		PFV	perfective
IMP	imperative		POSS	possessive
IPFV	imperfective		SG	singular
M	masculine			

References

Al-Ageel, H. (2015). Requests in Saudi Pidgin Arabic. *Business Management and Strategy* 6: 111–137. https://doi.org/10.5296/bms.v6i1.7682

Al-Azraqi, M. (2010). Pidginization in the Eastern region of Saudi Arabia: Media presentation. In Reem Bassiouney (ed.), *Arabic and the media. Linguistic analyses and applications*. Leiden: Brill. 159–174.

Albakrawi, H. (2012). The linguistic effect of foreign Asian workers on the Arabic Pidgin in Saudi Arabia. *Research on Humanities and Social Sciences* 2(9): 127–133.

Al-Bargi, A. (2014). Foreign accent determines local attitude. http://www.arabnews.com/news/567061.

Alghamdi, E. A. (2014). Gulf Arabic Pidgin: A descriptive and statistical analysis of stability. *International Journal of Linguistics* 6(6): 110–127. https://doi.org/10.5296/ijl.v6i6.6846

Al Hameli, A. (2013). Emirati life: Broken Arabic is unfair to language and learner. *The National*, (September 24, 2013). http://www.thenational.ae/thenationalconversation/lifestyle-comment/.

Al-Haq, F. A. and I. Al-Salman (2014). Jordanian Bengali Pidgin Arabic. *US-China Foreign Language* 12(5): 331–348.

Almoaily, M. (2008). *A data-based description of Urdu Pidgin Arabic*. MA dissertation, Newcastle University.

Almoaily, M. (2013). *Language variation in Gulf Pidgin Arabic*. PhD dissertation, Newcastle University.

Almoaily, M. (2014). Language variation in Gulf Pidgin Arabic. In I. Buchstaller, A. Holmberg, and M. Almoaily (eds.), *Pidgins and creoles beyond Africa-Europe encounters*. Amsterdam: John Benjamins. 57–83. https://doi.org/10.1075/cll.47.04alm

Al-Salman, I. A. K. (2013). *Jordanian Pidgin Arabic*. MA dissertation, Yarmouk University, Irbid.

Al-Sharkawi, M. (2005). *Arabicization: A case of second language acquisition*. PhD dissertation, University of Nijmegen.

Al-Sharkawi, M. (2007). Arabic Foreigner Talk. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich, A. Zaborski (eds.), *Encyclopedia of Arabic language and linguistics*, vol. 2. Leiden: Brill. 116–122.

Al-Sharkawi, M. (2010). *The ecology of Arabic. A study of Arabicization*. Leiden: Brill. https://doi.org/10.1163/ej.9789004186064.i-266

Alshuaimi, S. (2011). Why call it broken Arabic? http://suad47.blogspot.com/2011/07/why-call-it-broken-arabic.html. (20 August 2013)

Al-Shurafa, N. S. (2014). Linguistic functions of *fii/maafii* in Hejaz Gulf Pidgin Arabic. *King Khalid University Journal for Humanities* 21 (1): 11–27.

Al-Zubeiry, H. Y. A. (2014). Linguistic analysis of Saudi pidginized Arabic as produced by Asian foreign expatriates. *International Journal of Applied Linguistics & English Literature* 4 (2): 47–53.

Avram, A. A. (2010). An outline of Romanian Pidgin Arabic. *Journal of Language Contact* 3: 19–37. https://doi.org/10.1163/19552629-90000018

Avram, A. A. (2012). On the functions of *fi* in the verbal system of Arabic pidgins. *Romano-Arabica* 12: 35–58.

Avram, A. A. (2013). Multifunctionality in Gulf Pidgin Arabic: The case of *fi*. *Paper presented at "Linguistics in the Gulf 4", 17–18 March 2013*, Qatar University, Doha.

Avram, A. A. (2014). Immigrant workers and language formation: Gulf Pidgin Arabic. *Lengua y Migración* 6 (2): 7–40.

Avram, A. A. (2016). On the developmental stage of Gulf Pidgin Arabic. In G. Grigore and G. Biţună (eds.), Arabic varieties: Far and wide. Proceedings of the 11th international conference of AIDA, Bucharest, 2015. Bucharest: Editura Universităţii din Bucureşti. 87–97.

Avram, A. A. 2017. Sources of Gulf Pidgin Arabic features. In S. Bettega and F. Gasparini (eds.), *Linguistic studies in the Arabian Gulf, QuadRi – Quaderni di Ricognizioni*, vol. 7. Turin: Università degli Studi di Torino. 131–151.

Bakir, M. J. (2010). Notes on the verbal system of Gulf Pidgin Arabic. *Journal of Pidgin and Creole Languages* 25 (2): 201–228. https://doi.org/10.1075/jpcl.25.2.01bak

Bakir, M. J. (2014). The multifunctionality of *fii* in Gulf Pidgin Arabic. In S. Manfredi and M. Tosco (eds.), *Arabic-based Pidgins and Creoles, special issue of the Journal of Pidgin and Creole Languages* 29 (2): 416–442.

den Besten, H., P. Muysken, and N. Smith (1995). Theories focusing on the European input. In J. Arends, P. Muysken, and N. Smith (eds.), *Pidgins and creoles. An introduction*. Amsterdam: John Benjamins. 87–98.

Bizri, F. (2005). Le Pidgin Madam un nouveau pidgin arabe. *La Linguistique* 41 (2): 54–66. https://doi.org/10.3917/ling.412.0053

Bizri, F. (2009). Sinhala in contact with Arabic: The birth of a new pidgin in the Middle East. In R. Singh (ed.), *Annual Review of South Asian Languages and Linguistics*. Berlin and New York: Mouton de Gruyter. 135–148. https://doi.org/10.1515/9783110225600.133

Bizri, F. (2010). *Pidgin Madame. Une grammaire de la servitude*. Paris: Geuthner.

Bizri, F. (2012). Variation in verbal forms in Pidgin Madam and other Arabic pidgins. *Paper presented at Ninth Creolistics Workshop "Contact languages in a global context: Past and present", 11–13 April 2012*, Aarhus University, Aarhus.

Bizri, F. (2014). Unity and diversity across Asian migrant Arabic pidgins in the Middle East. *Journal of Pidgin and Creole Languages* 29 (2): 391–415. https://doi.org/10.1075/jpcl.29.2.07biz

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.

Brockett, A. A. (1985). *The spoken Arabic of Khabūra on the Batina of Oman*. Manchester: University of Manchester.

Dashti, A. A. (2013). Interacting with domestic workers in Kuwait: Grammatical features of Foreigner Talk. A case study. *International Journal of the Sociology of Language* 224: 63–84.

Ferguson, C. A. (1971/1996). Absence of copula and the notion of simplicity: A study of normal speech, Baby Talk, Foreigner Talk and pidgins. In T. Huebner (ed.), *Charles A. Ferguson, Sociolinguistic perspectives. Papers on language in society, 1959 – 1994*. New York: Oxford University Press. 115–123.

Ferguson, C. A. and C. E. DeBose (1977). Simplified registers, broken language, and pidginization. In A. Valdman (ed.), *Pidgin and creole linguistics*. Bloomington: Indiana University Press. 99–125.

Haraty, N., A. Oueini, and R. Bahous (2007). Speaking to domestics in Lebanon: Power issues or misguided communication?. *Journal of Intercultural Communication* 14. http://www.imoni.se/nr14/haraty.htm.

Mühlhäusler, P. (1997). *Pidgin and creole linguistics*, expanded and revised edition. London: University of Westminster Press.

Naylor, H. (2008). The word on the street – same-same, but different. *The National*, October 9, 2008. http://www.thenational.ae/news/uae-news/the-word-on-the-street-same-same-but-different.

Næss, U. G. (2008). "Gulf Pidgin Arabic": Individual strategies or a structured variety? A study of some features of the linguistic behaviour of Asian migrants in the Gulf countries. MA dissertation, University of Oslo.

Potsdam, E. and M. Alanazi (2014). *Fi* in Gulf Pidgin Arabic. *Kansas Working Papers in Linguistics* 35: 9–29.

Salem, A. A. (2013). Linguistic features of Pidgin Arabic in Kuwait. *English Language Teaching* 6 (5): 105–110.

Sebba, M. (1997). *Contact languages. Pidgins and creoles*. London: Macmillan.

Smart, J. R. (1990). Pidginization in Gulf Arabic: A first report. *Anthropological Linguistics* 32 (1-2): 83–119.

Tosco, M. and S. Manfredi (2013). Pidgins and creoles. In Jonathan Owens (ed.), *The Oxford handbook of Arabic linguistics*. Oxford: Oxford University Press. 496–519.

Tweissi, A. I. (1990). 'Foreigner Talk': Evidence for the universality of language simplification. In M. Eid and J. McCarthy (eds.), *Perspectives on Arabic Linguistics: Papers from the second annual symposium on Arabic linguistics*. Amsterdam: John Benjamins. 296–326. https://doi.org/10.1075/cilt.72.16twe

Versteegh, K. (2014a). Pidgin verbs: Infinitives or imperatives? In I. Buchstaller, A. Holmberg, and M. Almoailly (eds.), *Pidgins and creoles beyond Africa-Europe encounters*. Amsterdam: John Benjamins. 141–170. https://doi.org/10.1075/cll.47.07ver

Versteegh, K. (2014b). Speaking of the past. The development of past time reference in Arabic pidgins. In S. Manfredi and M. Tosco (eds.), *Arabic-based Pidgins and Creoles*, special issue of the *Journal of Pidgin and Creole Languages* 29 (2): 211–231. https://doi.org/10.1075/jpcl.29.2.02ver

Wiswall, A. (2002). Gulf Pidgin: An expanded analysis. http://oak.cats.ohiou.edu/~aw321500. (1 December 2002)

Zacharias, A. (2010) New lingo sets tongue wagging. *The National*, May 2, 2010. http://www.thenational.ae/uae-news/new-lingo-sets-tongue-wagging.

Mountains do not meet, but men do

Music and sociocultural networks among Arabic creole-speaking communities across East Africa

Shuichiro Nakao
Osaka University

Linguists have long assumed that Juba Arabic and Nubi, the two Arabic creoles spoken in East Africa, have been cut off from each other since their "linguistic divergence" in the 1880s. This historical interpretation, however, overlooks sociocultural (including linguistic) interactions between the Nubi-speaking communities of Uganda and Kenya and a minor Juba Arabic-speaking community in South Sudan called Malakiyyans since the 1880s down to the present day. This paper aims at exploring their history and the way in which they have interacted with each other to redefine their identity, focusing on the musical tradition called *dolúka* and *dirêr*.

Keywords: Nubi, Juba Arabic, creole, ethnomusicology, transnationality

1. Introduction: Questioning the linguistic divergence of Arabic creoles

There are two well-known Arabic creoles in East Africa, which have dialectal differences including some grammatical and lexical ones: Juba Arabic and Nubi (sometimes referred to as Kinubi, a Swahili term).[1] Juba Arabic is the main vehicular language of the multi-ethnic South Sudan and, at the same time, the native language of a large part of the urban population. Nubi is the vernacular language of the Nubi, a minority Muslim community of South Sudanese origins living in Ugandan and Kenyan towns and suburbs (mainly in Bombo, 21 miles north of Kampala, Uganda, and Kibera, southwest of Nairobi, Kenya). Nubi is also spoken by non-Nubi Muslim communities in northwestern Uganda (e.g., from Lugbara, Ma'di, Kakwa, and Alur

---

1. This study transcribes Nubi and Juba Arabic according to Nakao (2013a), since they basically share the phonetic/phonological system (except, for example, that /j/ is realized usually as plosive [ɟ] in Juba Arabic but affricate [dʒ] in Nubi).

ethnic groups) as one of their first languages. Some linguistic studies have noticed dialect divergence(s) within Nubi, although there has been no systematic study by sampling dialect-speakers born and raised in one place.[2]

The linguistic uniformity of Juba Arabic and Kinubi has been explained in the light of their common sociohistorical origin. It is traditionally argued that an Arabic pidgin had been stabilized by late 19th-century South Sudan,[3] before the main ancestors of the Nubi people left South Sudan for Uganda. For example, Nhial (1975: p. 81), in the earliest linguistic study on Nubi and Juba Arabic, argues:

> Some Juba Arabic speakers believe that it is an offshoot of Ki-Nubi. Linguistic similarities might seem to support this, but other factors tend to contradict it. For one thing, since the Nubis first settled in Uganda, there has been little contact between Uganda and the Sudan. For another, the centres where these varieties are spoken are not adjacent, but quite far from the common border. Beginning in the early 1960s, there has been some superficial contact, but this is obviously too recent to have resulted in a linguistic influence of one group upon the other. As such, the argument that Juba Arabic grew from Ki-Nubi can only be regarded as unfounded. It seems more likely that they both evolved from the military Arabic of the southern Sudan in the nineteenth century.

A number of studies also support the idea according to which speakers of Juba Arabic and Nubi have had few contacts since their displacement (Owens 1997: pp. 135–136, 160; Miller 2002: p. 22; Tosco & Manfredi 2013: pp. 503–504). In the same manner, it has been repeatedly argued that Ugandan and Kenyan Nubis have been isolated from each other for a long time.

This understanding of the history of Nubi and Juba Arabic, however, disregards the most important points. First, the present border between Uganda and (South)

---

2.   See Owens (1997: p. 160) and Tosco & Manfredi (2013: pp. 503–504). The lexical differences between Ugandan and Kenyan Nubi as of 1972–1975 given by Owens can no longer be observed, perhaps due to the refugee migrations from Uganda to Kenya (as Owens himself suggested). However, there are some easily recognizable morphosyntactic differences, although it would rather be a matter of degree. For example, the negative marker *mâ* often occurs in postverbal phrase position by central Ugandan Nubi speakers (Bombo, Kampala, Entebbe), whereas it often occurs in preverbal position by others (e.g., in Kenya and northern Uganda). Likewise, the sentence-final particle *ke* (indicating a polite emphasis) is widely used in Uganda but not much in Kenya. On the other hand, in Juba Arabic, the negator *mâ* always occurs in preverbal position and the particle *ke* is widely used. Furthermore, Nubi speakers are often well aware that their (or their distant friends' or relatives') own local Nubi variety borrows *more* words from dominant local languages that they are used to speaking; Dholuo in Kisumu, Lugbara in Arua, Luganda in Kampala, etc.

3.   In this article, I intentionally use the anachronic country names (such as South Sudan) to avoid geographic confusion.

Sudan was rectified in 1914, and Nubi settlements existed in what is today South Sudan in the early 20th century (Nakao 2016). Second, following the border rectification, some Nubis remained in South Sudan, and their descendants are now part of so-called Malakiyyans, a group of native speakers of Juba Arabic, living in urban quarters in South Sudan (Nakao 2013b). Third, in spite of their geographical separation, the Nubi and the Malakiyyan communities have maintained and developed sociocultural ties across the borders until the present day.

The aim of this paper is to redefine the history of Nubis and Malakiyyans from a sociocultural (including linguistic) viewpoint. Section 2 sketches the history of the linguistic and cultural creolization among the people in South Sudan, their subsequent migrations, and their sociocultural ties and interactions all across East Africa. Section 3 focuses on a musical tradition shared by Nubis and Malakiyyans (called *dolúka* by Nubis and *dirêr* by Malakiyyans) and shows how it has enabled them to reinforce their identity.

2. History of the Arabic creole-speaking communities in East Africa

2.1 Military slavery and musical traditions in 19th-century Sudan

It is widely accepted that the common ancestor of Juba Arabic and Nubi emerged during the 19th century, when Arabic was first introduced into South Sudan as a result of the Turco-Egyptian expansion under the Ottoman viceroy Muḥammad ʿAlī (Owens 1997).

In 1820, Muḥammad ʿAlī started to establish his "new model army" (*al-niẓām al-jadīd*), whose soldiers were drawn from black slaves who had been raided in what are today the border areas of Sudan and South Sudan. They were deployed in Egypt and the new Turco-Egyptian territories of Sudan where they became known as *jihādiyya* 'regular troops'. Most likely, these black slave soldiers had developed a reduced variety of Arabic. Casati (1891: p. 21) records an "Arabic" song sung by these soldiers, who had just returned to Sudan from the Syrian campaign in 1841 (transcribed according to the source text):

(1) *Ya tamra tamereteni,* O fruit, O fruit (my sweetness),
  Ya bent konti feni,   Where have you been, my girl?
  Kont and el ghendi,   I was with a gentle(man),
  Bakol kalava kendi,   Eating Indian sweets,
  Be nar el habib,   With the fire of the beloved,
  Ya abu Ibrahim.   O Father of Ibrahim.

This variety can be considered Egyptian Arabic except the form *feni* (i.e., Egyptian *fēn* 'where') in the first line, which rhymes with *tamereteni* (i.e., *tamratēnī* 'my two dates'), like *kendi* (i.e., *hindī* 'Indian') and *ghendi* (probably *gindī* 'soldier') do. Such an instance of paragoge, however, has not been attested in Egyptian Arabic, while it is quite common in Nubi (e.g., *wên ~ wéni* 'where'). Although the available data do not allow us to fully verify the linguistic nature of the Arabic variety spoken by these *jihādiyya* soldiers, it is clear from this instance that they did not only speak it but also sang in it at this stage.

In mid-19th century Khartoum, the military band of the *jihādiyya* army sang Arabic march songs, while black slave girls (also from around South Sudan) sang romantic songs in Sudanese Arabic at governors' parties (Sikainga 2010: p. 162).[4] These musical experiences influenced the creation of the modern Sudanese musical traditions, such as religious songs of the *zār-ṭumbura* spirit possession cult (Makris 2000: pp. 227–259) and non-religious songs sung in wedding ceremonies accompanied by the beats of goblet drums called *dallūka* (Malik 2003). Similar cultural traditions of South Sudanese ex-slaves have been also documented in Egypt (Walz 2012).

2.2   History and traditions of Nubis and Malakiyyans

In 1841, an expedition sent by Muḥammad ʿAlī reached Gondokoro (near Juba) in the Equatoria region, the southernmost part of South Sudan. Soon after, Equatoria was opened to ivory and slave traders from the north who, regardless of their country of origin (Sudan, Egypt, Malta, and other European countries), spoke Arabic as the trade language. Local populations, enslaved by these traders, learned Arabic from them and became interpreters at the trading posts. As Equatoria was annexed to Egypt in 1869, some *jihādiyya* regiments were sent into the region. In this context, new recruits were increasingly drawn from the local population, and interpreters were employed as native police-soldiers (Leonardi 2013: pp. 356–358). Given that both Nubi and Juba Arabic show substratum interference from local languages spoken in Equatoria, it seems that the ancestor of Nubi and Juba Arabic had crystallized in this sociolinguistic situation in mid-19th century (Owens 1997: pp. 160–163).

---

4.   More recently, Thorburn (1925) records Arabic marching songs of black Sudanese soldiers of the reformed Egyptian army (see 2.2), which resemble the *zār-ṭumbura* songs (Makris 2000: p. 233). Lopashich (1958) records Arabic dance songs of a Sudanese ex-slave woman who was brought to Montenegro by her husband around 1877.

In 1881, the Mahdī, Muḥammad-Aḥmad ibn ʿAbd Allāh, revolted against Turco-Egyptian rule in Sudan. As a result, the *jihādiyya* soldiers deployed in various areas of Sudan and South Sudan were either absorbed into the Mahdist army (also called *jihādiyya*) or took refuge in Egypt. By 1883, the Mahdī's revolt had wide-reaching effects in Equatoria. Many *jihādiyya* soldiers in the region gradually started to retreat southwards to central Africa (modern Uganda–Congo border areas) with their governor, Emin Pasha. In 1889, when Emin Pasha was finally "rescued" by Henry Stanley, a small part of these *jihādiyya* soldiers and their families followed Emin and Stanley to Tanzania to be "sent back" to Egypt or to be re-enlisted in the colonial German army in Tanzania.[5] Others were left behind in central Africa, until they met Frederick Lugard, who persuaded them to serve under the British in Uganda (including the easternmost parts of Equatoria) and Kenya (including the Jubaland, now southern Somalia). The ex-*jihādiyya* soldiers and their families who migrated to Uganda, Kenya and Tanzania became known as *Nubi*s. At that time they were stationed and settled near colonial posts such as Bombo and Kibera, Gondokoro (near Juba) and Nimule (on the South Sudan–Uganda border), before these districts were transferred to Sudan in 1914. Even after the border rectification, a few Nubi ex-soldiers remained in these towns (Nalder 1937: p. 60).

Meanwhile, in the 1880–1890s, another group of the ex-*jihādiyya* soldiers who had fled to Egypt or deserted from the Mahdist army was reorganized under the British leadership with the aim of fighting back the Mahdists to establish the Anglo-Egyptian rule over Sudan and South Sudan, whose southernmost post had been located at Mongalla (near Gondokoro) since 1901. Again, in Rejaf, there were still former *jihādiyya* soldiers, who remained in Equatoria during the Mahdist invasion. In 1927, when the Anglo-Egyptian colonial government decided to establish a new headquarters in Juba, the ex-soldiers were transferred from Mongalla, Rejaf and Gondokoro to be employed as civilian public workers. They were settled in a *malakiyya* 'civilian quarter' (*melekíya* in Juba Arabic), which was the first native lodging area in Juba. After the independence of Sudan in 1956, this neighborhood became the destination of rural migrants who accepted the Arabic variety of the ex-soldiers, now known as Juba Arabic, as the main urban lingua franca (Miller 2002: pp. 25–26; Leonardi 2013: pp. 359–360; Nakao 2013b, 2016).

---

5.   Interview with Mwalimu Salimu Kungulilo, Dar es Salaam, August 2014. He is a descendant of a member of this group. Nubis in Dar es Salaam had long lived in an area today called Unubini ('Nubi-land' in Kiswahili) in Chang'ombe (cf. Leslie 1963), but were evacuated from there under the recent urban development.

Reflecting the early history of military slavery, a large part of Nubi and Malakiyyan cultures has its base on the *jihādiyya* culture that had developed by the 1880s. For example, the *zār-ṭumbura* cult (called *túmbura* in Nubi and Juba Arabic) was practiced by both Nubis and Malakiyyans. However, today it is only practiced in South Sudan by those who have kinship ties with northern Sudanese practicing the same rite (Nakao 2013b).[6] One of the most important traditions shared by Nubis and Malakiyyans is the drum music mainly played in wedding ceremonies and called *dolúka* by Nubis and *dirêr* by Malakiyyans, which corresponds to the *dallūka* in northern Sudan. Perhaps one of the oldest songs sung by Nubis is the hand-game song in (2).[7] Interestingly, a Sudanese Arabic version of this song is recorded as a lullaby by Hillelson (1918), who said it originated in Egypt but by then was also prevalent in Sudan.

(2) *dawíya, dawíya, dawíya*   Dawiya, Dawiya, Dawiya
     wediní káki wâi   Give me a khaki
     káki fi jwo sondû   The khaki is in the box
     sondû m(â)=éndi muftâ   The box lacks the key
     muftâ fí na sultân   The chief has the key
     sultân ázu banâ   The chief wants girls
     banâ ázu mendîl   The girls want handkerchiefs
     mendîl fi dukân   The handkerchiefs are [sold] in the shop
     dukân ázu lében   The shop wants milk
     lében tete bágara   The milk is under the cow
     bágara ázu gési   The cow wants grass
     gési ázu mátara   The grass needs rain
     mátara wága, wawawa!   The rain fell, wa-wa-wah!

---

6.   Interview with Mwalimu Salimu Kungulilo in Dar es Salaam, August 2014. See also Meldon (1908), Leslie (1963: p. 48), Clark (1972: pp. 69–73) and de Smedt (2011: pp. 125–126) for Nubi "superstitions" including *túmbura*.

7.   By courtesy of Yakub Hassan and Ibrahim Harun in Kibera, August 2014. Mustapher Khamisy (in Bombo, August 2014) provided me with another version of this song, ending in *bágara ázu asîs / asîs fi jubâl / jubâl ázu mátara / mátara wága, wa!* "The cow wants grass / The grass is on the mountains / The mountains want rain / The rain fell, wah!" The same song with slight modifications is also known in Juba.

Figure 1. A Nubi hand-game song

2.3 Social interactions of Nubis and Malakiyyans across East Africa

Despite their dispersion throughout northeastern Africa, the ex-*jihādiyya* popula-
tions have not been completely isolated from each other. As early as in the 1890s,
Nubi communities in the British and German spheres of East Africa were joined
by their former comrades-in-arms who fled to Egypt (Meldon 1908; Moyd 2014).
Sometimes Nubis on the British and German sides happened to fight against each
other. Nubi oral traditions in Kibera and Arua report that during World War I at
the warfront in Tanga (Tanzania), Nubi soldiers under the British heard their en-
emies issuing commands in Arabic, "*ádarab nútfa! ádarab nútfa!* (Fire the canon!
Fire the canon!)".[8] Once they realized that their enemies were actually their kins,
they stopped fighting. After World War I, some of these Nubis who formerly served
under the Germans were incorporated into the King's African Rifles under the
British (Johnson 2009: p. 117).

---

8. This does not seem to be Nubi (cf. Nubi. *dúrubu mútufa* "fire the canon"). Actually, a minority
of early Nubis who were recruited in Egypt and in northern Sudan spoke non-creole Egyptian
Arabic in the early 20th century (Raddatz 1892; de Smedt 2011: p. 120; Moyd 2014: p. 96; Nakao
2016; see also Sikainga 2010: p. 163). Today, virtually no Tanzanian Nubi speaks Nubi as the
vernacular language, unless s/he has lived among Kenyan or Ugandan Nubis.

During and after the World War II, Nubis and Malakiyyans (as well as other Nubi-speaking Ugandans and Juba Arabic-speaking South Sudanese) often crossed the borders to look for their relatives and/or to find military employment in another country (Leopold 2006: p. 193; Johnson 2009: pp. 117–118; de Smedt 2011: pp. 123–124). Such cross-border migrations were accelerated by the outbreak of civil wars in pre- and post-independence Sudan (1955–1972, 1983–2005, as well as during the ongoing civil war in South Sudan since 2013) and the fall of the Ugandan President Idi Amin (a Nubi-speaking Kakwa) in 1979, producing many Nubi- and Juba Arabic-speaking refugees.[9]

As a result, Nubi and Malakiyyan communities have kept strong ties across political borders. Among Ugandan and Kenyan Nubi communities, according to Clark (1972: pp. 201–204), not only did the kinship system entail reciprocal visits for their relatives' weddings, funerals or serious illness, but Nubi boys who finished schooling would also spend a certain period of time with their kin in another country, and cross-border recruitment through the influence of kinship ties has been quite common. Another prominent example is Abdel Rahman Sule, one of the earliest South Sudanese politicians (for details, see Nakao 2013b). During his life, he often crossed political borders to enhance South Sudanese-Ugandan relations by making use of Malakiyyan-Nubi ties (Kuyok 2015: pp. 100–103).

As such, it is sometimes difficult to tell if a person is a Nubi or a Malakiyyan. For example, Yusuf Fataki (1939–2004), the first South Sudanese artist to sing in Juba Arabic on air, is remembered in South Sudan as a "Malakiyyan" (or a member of the Kakwa ethnic group) born in Yei, South Sudan, but in Uganda, he is remembered as a "Nubi" born in Arua (Uganda). According to his brothers in Arua, their father Khalifa Fataki Saghir was a Kakwa from Gulumbi (near Yei) who served under both Emin Pasha and the British, and played a role in the spread of Islam in Uganda, an archetypal Nubi figure. His 56 children are spread all across Sudan, South Sudan, Uganda, Kenya and Tanzania. One of these children was Yusuf, who, just after the

---

9. These migrations actually induced new linguistic contacts between Nubi and Juba Arabic speakers. The Nubi data recorded by Nhial (1975) were furnished by a South Sudanese, who "spent 9 years as a refugee in Uganda, much of that time in Nubi homes" (Ḥurreiz & Bell 1975: p. 80). The speaker observed that Ugandan Nubis were interested in learning Juba Arabic from South Sudanese refugees so they could replace Swahili or Luganda loanwords in Nubi in their speech (Nhial 1975: p. 92). Ustaz Mustapher Khamisy in Bombo, a famous grass-roots Nubi writer, explains that he became passionate about regaining "original Nubi (*núbi taasíli*)," eliminating Swahili or Luganda influence when he encountered Juba Arabic while he was in exile in Yei, South Sudan (interview in Bombo, August 2014). Clark (1972: pp. 282–283) mentions that South Sudanese refugees were integrated into the social activities of the Kibera Nubi community due to their linguistic affinity, and Labidi (2014) notes that Nubi speakers benefitted from their knowledge of Nubi during their exile in South Sudan in seeking business opportunities.

outbreak of the first Sudanese Civil War in 1955, was in a Sudanese prison where he would sing sweet songs he had learned or composed as a *dolúka* and *dirêr* singer. One of these songs was *Yei Beledina*, for which he was released and later promoted as a "patriotic" singer:[10]

(3) *yêi belédina*   Yei, our country
     wa kúlluna ikhwân   and we are all brothers
     sudáni belédina, yêi   Sudan is our country, Yei
     ya akwána ámsuku   O brothers hold
     béled kuwési ya   [our] country strongly
     ya jáma kelína kúruju   O people let us cultivate
     lúbiya fi belédina   beans in our country

It is interesting to note that, although Yusuf retired from singing in 1963, his influence has lasted until today among South Sudanese artists. According to Lorins (2007: p. 181), Derik Alfred, the managing director of Kwoto Culture Centre, which has long been famous for representing "South(ern) Sudanese identity" in Khartoum by performing songs and dramas in Juba Arabic (as well as in South Sudanese vernaculars) since 1994, related in a 2002 lecture that they considered Malakiyyans represented by Yusuf Fataki as precursors of their activity.[11] In other words, one of the roots of modern Juba Arabic popular culture is sought in the transnational Nubi-Malakiyyan culture.

Many historical studies have shown that the social network of Nubis in Uganda, Kenya and Tanzania has been exercised more prominently in political domains (Leslie 1963: p. 48; Clark 1972; Parsons 1997; Johnson 2009; de Smedt 2011). Various Nubi voluntary associations, starting around the 1920–1930s, became active in the 1940s to reinforce the "migrant Sudanese" identity by uniting the scattered Nubi communities in reaction to the colonial government's attempts to integrate them into local "tribal" administration. For example, the Kenya, Uganda and Tanganyika branches of the Sudanese Association of East Africa, whose goal was to "unite the many isolated Sudanese communities scattered throughout East Africa to press the British government to grant them permanent title to their land" (Parsons 1997: p. 109), held a grand meeting at Bombo in 1948 (Clark 1972: p. 94).

---

10. Interview with Abdalla Fataki and Zakariya Fataki in Arua, September 2015. Even today, this song is performed by younger South Sudanese artists, substituting "Sudan" with "South Sudan" (https://www.youtube.com/watch?v=ZoEbt7CGhG8, accessed on February 23, 2016).

11. In 2012, Derik Alfred, along with Joseph Abuk, who founded Skylarks Dramatists' Association in 1979, directed *"Cymbeline"* in Juba Arabic on the stage of the World Shakespeare Festival in London.

Photo 1. Sudanese Association of East Africa, Nairobi Branch (Kibera)[12]

Photo 2. Sudanese Association of East Africa, Bombo Branch[13]

---

12. By courtesy of Abuba Fatma, Kibera, August 2014.

13. By courtesy of Mzee Ismail Abderrahman Rehan Dabule, Kampala, August 2014.

Such voluntary associations claiming a "stateless" and "migrant" identity were, in the end, demolished by next generations of Nubis in the political process of integrating themselves into their local states after independence (Parsons 1997; Johnson 2009; de Smedt 2011).[14] After long struggles, Nubis managed to acquire official recognition as a "native ethnic group" in Uganda and Kenya (in the 1995 constitution of Uganda and the 2009 census in Kenya), but in spite of these efforts they have remained marginalized until the present day, and, to their sorrow, the Nubi language and traditions are diminishing among the younger generations, even in Kibera (Constantine 2011).

In response to the present situation, Nubis in Kenya and Uganda have recently started to redefine and promote their identity as "transnational" (rather than "migrant" and "stateless") minority through cultural activities, represented by an annual international cultural festival popularly called *chai* 'tea [party]' (in Swahili).[15] Interestingly, Malakiyyans in post-independence South Sudan, where they, as a mainly Muslim community, have become a religious minority, are starting to join in *chai*.

3.   Performing identity through *dolúka* and *dirêr*

3.1   Organization of *dolúka* and *dirêr*

As mentioned earlier, *dolúka* and *dirêr* music have been strongly related to the Nubi and Malakiyyan wedding traditions, which play a symbolic role to their kinship ties and social network. As such, *dolúka* and *dirêr* have also been performed on many occasions apart from wedding ceremonies, such as for Saturday night entertainment, communal ceremonies and, of all things, the *chai* festival.

Actually, in the past, *dolúka* was mainly provided by specific women's associations, which, in Kenya, together with male-dominated football clubs, and was strongly associated with the Nubi voluntary associations (Clark 1972: pp. 66–69). According to Clark (1972: pp. 87–96), inspired by the 1948 meeting of the Sudanese Association, such football clubs and *dolúka* clubs started to organize exchange visits between Kenya and Uganda in the late 1950s. Today, as Nubis have no official international communal organization, *chai* functions as an important occasion for Nubi

---

14. For example, later, the Bombo branch of the Sudanese Association of East Africa was reorganized, with the new name "Nakatonya Islamic Community" so as to avoid the word "Sudanese."

15. Apart from *chai*, for example in Kibera, two "Nubian Cultural Ambassadors" were unveiled in a Nubi cultural contest in 2013. Yakub Hassan in Kibera, who had set up a Facebook page for the Nubi language in 2011, is currently compiling a Nubi dictionary.

communities to communicate with each other through cultural activities (*dolúka*, football matches, craft exhibition, keynote speeches, home visits and so on).

Chai is organized on both Ugandan and Kenyan sides at the same time during the Easter season, and the host community is chosen on a rotational basis. For example, the festival was organized at Kibera and Bombo in 2015 and at Mombasa (Kenya) and Busia (Uganda) in 2016. On the Ugandan side, *chai* has taken place not only in such larger Nubi settlements, but also at historical sites, such as Ajulu, Boroli and Pajao, where Turco-Egyptian or British colonial stations were located, and even in South Sudan (Labidi 2014: p. 35). As such, in recent years, even South Sudanese *dirêr* clubs have begun to participate in *chai* on the Ugandan side, and they have even had their songs recorded at the Uganda Broadcasting Corporation and broadcasted on a Nubi radio program. On the Kenyan side, *chai* has been going through an exciting reformation. Empowered by the Grandpa Records company founded by Yusuf Noah (a.k.a. Refigah Heviweit), formerly a Nubi hip-hop artist, it is organized as an international spectacle that "will see members of the Nubian [sic] community come from every corner of this region, all the way from [South] Sudan, Uganda, Kenya and Tanzania."[16]

Picture 3. *dolúka* at a wedding, photographed by the author at Bombo, 2014 (Yal Hamam)

---

16. "Kenya: Grandpa Records to Relaunch the Nubian Cultural Nite", *AllAfrica*, February 19, 2015 (http://allafrica.com/stories/201502191366.html, accessed on February 21, 2016). One of the prominent musicians of Grandpa Records, Alfatih Philip Abbas Kabush (a.k.a. Unique Arafat) is from Sudan but grew up among Nubi communities. "Descendants of Africa's Nubian tribe keep culture alive through hip-hop", *AP*, May 9, 2006 (http://www.aparchive.com, accessed on February 21, 2016). His father is a veteran Nuba politician.

Picture 4. *dirêr* at a wedding, photographed by the author at Juba, 2013 (Al-Aman Social Club)

Despite the difference in name, *dolúka* and *dirêr* are almost the same culture except for minor terminological differences.[17] Basically, a *dolúka* and *dirêr* performance begins around sunset and continues throughout the night; after tightening the drum heads with fire, the performers begin by making regular ticks by beating a *kóngkóng* mini-drum with a stick, followed by polyrhythmic beats of two or three middle-sized drums beaten with hands (called *fádulu-kénya* 'remain [in] Kenya' by Nubis and *sévén-kéya* 'seventh KAR [King's African Rifles]' by Malakiyyans), two small drums beaten with sticks (called *kálif* 'roll' by Nubis and *kelîf* 'rolling' by Malakiyyans), and a large 'mother drum' beaten with hands (called *uma-lungára* by Nubis and *uma-nugára* by Malakiyyans), accompanied by a pair of rattles made of cans and sand (called *koyo* by Nubis and *koskôs* by Malakiyyans).[18]

The music is provided by a local *dolúka* or *dirêr* club: For example, there are Kibera Social and Sports Club, Sister Club, Yal Safina and others in Kibera, Bombo Social and Sports Club, Sister Club, Yal Hamam and others in Bombo, and each Ugandan Nubi community in Entebbe, Arua and Masindi has its own Amani Social Club. In Juba, today, there is a single *dirêr* club called Al-Aman Social Club. A *dolúka* or *dirêr* club consists of both female and male members, and the female

17. Many Malakiyyans consider *dirêr* to be a culture descended from Mongalla, and they do not directly connect it with Nubi (although that account is not very probable, since Mongalla is too new). Etymologically, *dirêr* is related to *deríra*, a head cloth worn in wedding ceremonies.

18. Interview with members of Yal Hamam in Bombo, September 2014, and Al-Aman Social Club in Juba, October 2013.

members wear traditional clothes called *gurbába* as a uniform of the club. Although space does not permit further details here, while the music is performed, audiences make lines according to gender to perform simple dance steps.

3.2   Lyrics and languages of *dolúka* and *dirêr* songs

Backgrounded by the common early history and transnational interactions, many old songs sung in *dolúka* and *dirêr* are shared by all Nubi and Malakiyyan communities (except Tanzanian Nubis who sing in Swahili), while new songs are continuously composed. The main themes of the lyrics are a celebration for marriage and provocative gossip (as so are Sudanese *dallūka*; see Malik 2003), sometimes mixed with military memories and communal identity. Most *dolúka* and *dirêr* songs are composed in Nubi or Juba Arabic, but in general, its linguistic register is often slightly different from the spoken variety; it sometimes exhibits fossilized non-creole Arabic forms, mixtures of Nubi and Juba Arabic forms and irregular borrowings from African languages. As a result, the lyrics of older songs can sometimes be obscure.

Let us first compare versions of the most famous tune *Abu Jarara* 'The Buttoned One' from Juba (4), Kibera (5) and Arua (6).[19]

(4)   *abu jarára*   The buttoned one,
      jararú le wiláya   buttoned up (?) to the province
      lel banât ta sudân for the girls of Sudan
      nenzilí yôm sába   Grounded for seven days
      ma rijâl fi midân   with men in the square

Figure 2. "Abu Jarara" (as sung in Juba)

19. By courtesy of the members of Al-Aman Social Club in Juba, October 2013; Abuba Segiya in Kibera, August 2014; Abdul-Juma Labidi in Arua, September 2015 (Labidi 2014: p. 34). "Buttoned" may mean "well-dressed" (Rombek Logworong, p.c.).

(5) *abu jarára*   The buttoned one
     nenzilú lel uláya   was put down for Britain
     lel banâ ta sudân   for the girls of Sudan
     wúsule yôm sába   Seven days passed,
     mal-rijâl fi midân   the men's bride price is in the square

(6) *abu jarára jarára*   Those who buttoned themselves
     min uláya   [came] from Britain
     yal-banîn ta sudân   for the Sudanese youth.
     gedimú yôm sába   They (the British) took [and punished]
     ma rujâl fi midân   them in the field for seven days

The singers interpret the meaning of the song variously, for example, as celebrating the wedding ceremony or remembering the colonial days of military service. Linguistically, it is interesting to note that, in the form *lel* 'to/for the' appearing in the versions of Juba and Kibera here, the Arabic definite article (nonexistent in Nubi and basilectal Juba Arabic) is fossilized.[20]

The next *dolúka* song from Arua depicts the Nubi's war experience during World War I serving under the British.[21] In this song, a non-creole form, *gelbí* 'my heart' is found (compare Nubi *gélba taí* and Sudanese Arabic *gaḷbī*); this conforms to the attestations of the non-creole variety of Arabic spoken by a group of Nubis around the early 1910s (Nakao 2016).

(7) *hukum-nasára wája gelbí*   The British rule pained my heart
     ya akwána   O my brothers
     ína dúsman ma jéremani   We fought the Germans
     ína rásulu fi jebel-rwánda   As we reached Mt. Rwanda (?)
     kamân ma taliyân   Again, [to fight] with Italians

At times, military experience is metaphorically related to marriage. The next *dirêr* song from Juba remembers the farewell ceremony for their comrade Bangladeshi troops in World War II.[22] The comradeship is, then, used as a metaphor to encourage the new relation by marriage.

---

20. Compare Sudanese Arabic *le* = *l* = *banāt* (for = DEF = girl.PL) 'for the girls' with Nubi *na banâ* (for girl.PL) and basilectal Juba Arabic *le banât* (for girl.PL) 'for (the) girls'.

21. By courtesy of Abdul-Juma Labidi, Arua, September 2015 (see also Labidi 2014: p. 34).

22. Interview with the members of Al-Aman Social Club, Juba, October 2013.

(8) *aju-bádu kuwési*   Liking each other is good
     kelí nína límu bádu   Let us get together
     bangaladési gí rúwa   Bangladeshis are departing
     kelí nína rúsu móyo   Let us sprinkle water [for celebration]
     nas-arúsa fí wéni   Where are the bride, her family and friends?
     kedé úmon tála bára   Let them come out
     bangaladési gí rúwa   Bangladeshis are departing
     kelí nína límu bádu   Let us get together

As mentioned above, gossip is another important theme for *dolúka* and *dirêr*. The next song from Kibera is allegedly known as one of the oldest *dolúka* songs. It is difficult to tell what was the original theme(s) of this song was (were), but the second stanza seems to show that the composer intended to respond to a gossipy *dolúka* song in which she was ridiculed.[23] From a linguistic viewpoint, it is interesting to note that the lyrics include the completely cryptic word *tambéle* ('remote country?') and the Luganda word *wééraba* 'good-bye', which is almost cryptic to most Kenyan Nubis as well.

(9) *ibe deíya tambéle*   Ibn Daḥiyya, *tambele* (?)
     aíya ibe deíya   *aiya*, Ibn Daḥiyya
     ibe deíya wééraba   Ibn Daḥiyya, *good-bye*
     sála wonusú ána je íja   Although I was the subject of gossip
     ána miskíni   I am poor
     mâ bu logó hája min ána   Nothing will be taken from me

The next gossip songs are relatively new *dirêr* songs from Juba. As seen in these examples, the lyrics of *dirêr* songs sometimes include Nubi and Swahili words.[24] The song (10) begins with Swahili words, *hodi* 'excuse me' and *karibu* 'welcome', to introduce a secret meeting at night. As well, in the song (11) reprimanding a singer who cheated on another singer, in addition to Swahili *haraka* 'quickly', Nubi forms *na* (dative preposition) and *rági* 'man' are used along with corresponding Juba Arabic forms *le* and *rágil*, and Nubi negative construction *ána áju mâ* (1SG want NEG) "I don't want" (contrastive to Juba Arabic word order *ána mâ áju*) appears. Moreover, *gidída* 'chicken' in (10) and *akilí* 'to be fed' in (11) are also Nubi forms corresponding to *gidéda ~ gidáda* and *akilú* in Juba Arabic.

---

23. Interview with Abuba Segiya, Kibera, August 2014. Ibe Deiya is remembered as one of Emin's soldiers from Darfur (Labidi 2014: p. 16). According to Abuba Segiya, *tambéle* means 'somewhere far away'. Interestingly, Tambili (in South Sudan) located just south to Darfur is (allegedly) found in a northern Sudanese *ṭumbura* song (Makris 2000: p. 263).

24. By courtesy of the members of Al-Aman Social Club, Juba, October 2013.

(10) *kóng kóng, hódi*   Knock-knock, excuse me
     karíbu   Yes please
     munú dúgu bâb dé   Who's knocking on the door?
     dé ána, halû   It is I, hello
     dé ána ásuma góho   It is I who heard a cough
     galí gidída taláta, mamá   Saying that, three hens, mommy,
     áse dé ána dába le halû   I slaughtered them for Mr. Hello

(11) *halíma wé, kélim le dafála*   O Ḥalīma, tell Dafʿallāh
     kedé úwo jíbu sáa haráka   To bring a watch quickly
     dusuman-ganá,   Singers' conflict
     kélim na akwána   Tell brothers and sisters
     akilí rági ma dáwa   That a man was poisoned
     shulu-rágil ta mára,   Taking someone's husband,
     ána áju mâ   I don't want that
     kúlu wâi ma tô   Every one [should be] with her own

The last is one of the newest nationalistic *dirêr* songs from Juba, singing of the Comprehensive Peace Agreement in 2005, that brought "peace" to South Sudan.

(12) *yómna, yómna,*   Our day, our day,
     junubîn léna   Southerners, [come] to us
     aléla yómna   Today is our day
     baba-sálva rúwa fi nevásha   Baba Salva went to Naivasha
     rúwa jíbu salâm léna   And brought us peace
     aléla yómna   Today is our day
     mama-rebéka rúwa fi entébe   Mama Rebecca went to Entebbe
     rúwa jíbu salâm léna   And brought us peace
     aléla yómna   Today is our day
     wíhida wataníya, wíhida   National unity, unity

Some Kibera Nubis told me that they had heard a song in a recent *chai* in Kibera in which "the name of the president of South Sudan was mentioned" and that "old Nubi words like *aléla* 'today' (instead of Nubi *naáré*) were used". Although it is unclear if it was this very song, and why it was performed among Nubis in Kibera, we can at least conclude that the sociopolitical situation of South Sudan, the "homeland" of the Nubis, was certainly on their minds.

4.   Conclusion

This paper has explored the more than a century-long history of kinship ties and sociocultural interactions between Nubi and Malakiyyan (as well as other Nubi or Juba Arabic-speaking) communities scattered throughout East Africa, in which they have developed the transnational identity. As shown in the last section, the *dolúka* and *dirêr* music has played a significant role in this process, and their lyrics exhibit something of the complicated history of their interactions, although they would be too sporadic and ambiguous to be evidence for the linguistic convergence of Nubi and Juba Arabic.

On the last day of my short fieldtrip in Arua, September 2015, Mustafa, a Nubi gentleman who kindly took care of me, and Yusuf, a Nubi-speaking Lugbara, with whom I had made friends in Juba in 2009, bade me farewell with a Nubi proverb: *jubâl mâ gí límu, lakín binádam bi já límu ma binádam* "Mountains do not meet, but a man may sometime meet [another] man". As this proverb states, in the history of Nubis and Malakiyyans, men actually did meet with each other again and again, to re(-)create themselves. Now that South Sudan has seceded from Sudan and entered the East African Community, Nubi and Juba Arabic may come to be even more and more in contact with each other.

Acknowledgment

References

Casati, G. (1891). *Ten Years in Equatoria and the Return with Emin Pasha* (translated from Italian by J. R. Clay and W. S. Landor). Vol. 1. London: Frederick Warne.

Clark, D. (1972). *Kibera: Social Dynamics of a Low Income Neighbourhood in Nairobi*. MA thesis: Makerere University.

Constantine, G. (2011). *Kenya's Nubians: Then & Now*. Bangkok: Nowhere People.

de Smedt, J. (2011). *The Nubis of Kibera: A Social History of the Nubins and Kibera Slums*. Ph.D. thesis, University of Leiden.

Hillelson, S. (1918). Arabic Nursery Rhymes. *Sudan Notes and Records* 1, 25–29.

Ḥurreiz, S. H. and H. Bell (eds.) (1975). *Directions in Sudanese Linguistics and Folklore*. Khartoum: Khartoum University Press.

Johnson, D. H. (2009). Tribe or Nationality? The Sudanese Diaspora and the Kenyan Nubis. *Journal of Eastern African Studies* 3 (1): 112–131. https://doi.org/10.1080/17531050802401825

Kuyok, K. A. (2015). *South Sudan: The Notable Firsts*. Bloomington: Author House.

Labidi, A. J. (no date, 2014). *Understanding the Nubi of Eastern Africa*. Privately published.

Leonardi, C. (2013). South Sudanese Arabic and the Negotiation of the Local State, c. 1849–2011. *Journal of African History* 54 (3): 351–371. https://doi.org/10.1017/S0021853713000741

Leopold, M. (2006). Legacies of Slavery in North-West Uganda: The Story of the "One-elevens". *Africa* 76 (2): 180–199. https://doi.org/10.3366/afr.2006.76.2.180

Leslie, J. A. K. (1963). *A Survey of Dar es Salaam*. London: Oxford University Press.

Lopashich, A. (1958). A Negro Community in Yugoslavia. *Man* 58: 169–173. https://doi.org/10.2307/2795853

Lorins, R. M. (2007). *Inheritance: Kinship and Performance of Sudanese Identities*. Ph.D. thesis, The University of Texas at Austin.

Makris, G. P. (2000). *Changing Masters: Spirit Possession and Identity Construction among Slave Descendants and Other Subordinates in the Sudan*. Evanston: Northwestern University Press.

Malik, S. I. (2003). *Exploring Aghani al-Banat: A Postcolonial Ethnographic Approach to Sudanese Women's Songs, Culture, and Performance*. Ph.D. thesis, Ohio University.

Meldon, J. A. (1908). Notes on Sudanese in Uganda. *Journal of the Royal African Society* 7 (26), 123–146.

Miller, Catherine. (2002). The Relevance of Arabic-based Pidgin/Creoles for Arabic Linguistics. In G. Mansour and M. Doss (eds.), *Contributions to Arabic Linguistics*. Cairo: Arab Research Centre. 7–45.

Moyd, M. R. (2014). *Violent Intermediaries: African Soldiers, Conquest, and Everyday Colonialism in German East Africa*. Athens: Ohio University Press.

Nakao, S. (2013a). The Prosody of Juba Arabic: Split Prosody, Morphophonology and Slang. In M. Lafkioui (ed.), *African Arabic: Approaches to Dialectology*. Berlin: de Gruyter. 95–120. https://doi.org/10.1515/9783110292343.95

Nakao, S. (2013b). A History from Below: Malakia in Juba, South Sudan, c. 1927–1954. *The Journal of Sophia Asian Studies* 31: 139–160.

Nakao, S. (2016). More on Early East African Arabic Pidgin. In George G. and G. Bițună (eds.), *Arabic Varieties: Far and Wide. Proceedings of the 11th International Conference of AIDA – Bucharest* 2015. Bucharest: Editura Universității din Bucureşti. 413–420.

Nalder, L. F. (ed.) (1937). *A Tribal Survey of Mongalla Province*. London: Oxford University Press.

Nhial, A. A. J. (1975). Ki-Nubi and Juba Arabic: A Comparative Study. In Ḥurreiz & Bell (pp. 81–93).

Owens, J. (1997). Arabic-based Pidgins and Creoles. In S. G. Thomason (ed.), *Contact Languages: A Wider Perspective*. Amsterdam: John Benjamins. 125–172. https://doi.org/10.1075/cll.17.08owe

Parsons, T. (1997). "Kibra is Our Blood": The Sudanese Military Legacy in Nairobi's Kibera Location, 1902–1968. *International Journal of African Historical Studies* 30 (1): 87–122. https://doi.org/10.2307/221547

Raddatz, H. (1892). *Die Suahili-Sprache, enthaltend Grammatik, Gespräche, Dialekte aus dem Innern und Wörterverzeichnisse, mit einem Anhange: Sudan-Arabisch, und einer Einführung in die Bantusprachen*. Leipzig: Koch's Verlag.

Sikainga, A. A. (2010). Slavery and Social Life in Nineteenth-Century Turco-Egyptian Khartoum. In T. Walz & K. M. Cuno (eds.), *Race and Slavery in the Middle East: Histories of Trans-Saharan Africans in Nineteenth-Century Egypt, Sudan, and the Ottoman Mediterranean*. Cairo: The American University in Cairo Press. 147–170.

Thorburn, D. H. (1925). Sudanese soldiers' songs. *Journal of the African Society* 24 (96): 314–321.

Tosco, M. & S. Manfredi. (2013). Pidgins and Creoles. In J. Owens (ed.), *The Oxford Handbook of Arabic Linguistics*. Oxford: Oxford University Press. 495–519.

Walz, T. (2012). Nubians, Blacks Africans and the Development of Afro-Egyptian Culture in Cairo at the End of the Nineteenth Century. *Sudan Studies Association Bulletin* 30 (1): 3–11.

Determiner phrase

How specific is it in Moroccan Arabic-French codeswitching?

Karima Ziamari

GRAL, Faculty of Letters and Human Sciences-Meknes, Morocco / LACNAD-Inalco, Paris

Nominal insertions in Moroccan Arabic-French codeswitching are very common. They typically appear as French maximal projections embedded in a larger constituent headed by the Arabic determiners *wāḥəd* and *hād*. However, the reasons behind the insertion of determiners have not been clarified. This study, which relies on the Matrix Language Frame model (Myers-Scotton 1993), seeks to elucidate the factors inducing the insertion of determiners in the morphosyntactic and semantic frame of Moroccan Arabic. Analyzing eleven hours of recorded data, we will show that on the morphosyntactic level, the mismatch between Moroccan Arabic and French definiteness, gender and number may explain the frequency of such insertions. Though, morphosyntactic structure is not the only factor at play in contexts where determination is complex in both languages, and we thus need to take into account other domains such as the semantic, pragmatic and enunciative ones.

Keywords: determiner phrase, Moroccan Arabic-French codeswitching, Moroccan Arabic, French, language contact, determiners, nominal insertion, MLF model

1. Introduction

Nominal insertion has been widely studied in many language pairs (Boumans 1995; Muysken 2000, 2008), and Moroccan Arabic-French codeswitching seems not to deviate from the norm (Ziamari 2008, 2012, 2013). However, when it comes to Moroccan Arabic-French contact, Myers-Scotton (2002) speaks about "*problematic data*" because of a number of idiosyncratic traits of the nominal insertion. In this context, the reasons laying behind the insertion of some determiners have not been

© 2018 John Benjamins Publishing Company

clarified yet. This paper aims at revisiting some structures involving determiners in this language pair and offering a new analysis of the insertion of determiners in the morphosyntactic frame of Moroccan Arabic as a ML. The study is based on the insertional model of codeswitching developed by Myers-Scotton (1993, 2002) and takes a stand against linear approaches that have extensively dominated the analysis of Moroccan Arabic/French codeswitching (Bentahila & Davies 1983; Lahlou 1991). The questions raised by the paper can be summed up as follows: why are French determiners so productive in Moroccan Arabic French codeswitching? Why do they particularly appear after some Moroccan Arabic determiners like *wāḥəd* and *hād*?

2. Informants and data

The study is based on 11 hours of recoded data produced by Moroccan students in two different communicative settings. The first one includes formal situations such as classrooms, while the second one gives evidence of informal interactions in students' campus, cybercafés, and cafés. My informants, who were my students at the time of the data collection, are fluent in both Moroccan Arabic and French. Arabic is their L1 (first language), whereas French is compulsory in their curriculum. Thirty-three informants (10 females and 23 males), aged between 18 and 30 participated in the investigation which was conducted over a period of three years (from October 1998 until June 2001).

3. The theoretical background: The MLF model

Myers-Scotton's Matrix Language Frame model (hereafter MLF) is a model of bilingual language production (Myers-Scotton 1993; Myers-Scotton & Jake 2015). There are two fundamental dichotomies in the MLF model: the Matrix Language (ML) vs. the Embedded Language (EL), and content morphemes vs. system morphemes. The ML typically sets the morphosyntactic frame of the bilingual CP (complementizer phrase), which represents the domain of analysis. This means that the ML provides the system morphemes for controlling the distribution of elements over the clause:

> The ML is the source of the abstract grammatical frame of the CS clause; the role of the EL is limited largely to supplying content elements and peripheral monolingual (EL) phrases […]. Our unit of analysis is the clause, or CP, the projection of complementizer, or COMP.   (Myers-Scotton & Jake 2015: p. 418)

4.   Nominal insertion in Moroccan Arabic-French codeswitching

One of the most striking features of this language pair resides in nominal insertion (Bentahila & Davies 1983, 1998, 2007; Bentahila et al. 2013; Chan 2009; Heath 1989; Lahlou 1991). If other language pairs have very rarely this type of insertion (see Backus & Ziamari 2006 for Turkish-Dutch codeswithing), Moroccan Arabic and typologically-related languages such as Algerian Arabic (Caubet 1998) and Tunisian Arabic (Poplack et al., 2015) in contact with French exhibit this feature frequently. The occurrence of determiners when one of these two languages is involved by codeswitching has not ceased to attract the attention of specialists (Aabi 1999; Boumans 1995, 1998; Boumans & Caubet 2000; Gardner-Chloros 2009; Muysken 2000). Myers Scotton (2002) refers to this context as "*problematic codeswiching data*" for assessing the relevance of the MLF in the description of these language pairs. Poplack et al. (2015: p. 177), on their part, state that "*language pairs involving Arabic often qualify that language as resistant to constraints found to operate elsewhere*". It is by analyzing such nominal structures in Arabic-French codeswitching that linear models were confronted with the peculiarity of this structure (Muysken 2000: p. 83) and thus introduced the concept of insertion of constituents resisting to constraints (Sankoff & Nait M'barek 1988). According to the MLF model, Moroccan Arabic-French codeswitching generates three different nominal structures: a NP, called mixed constituent, where Moroccan Arabic provides the determiner and French provides the noun (Example 1); internal EL islands formed of maximal projections, where the NP is supplied by French (Example 2) and full French NPs (Example 3).

(1)   *dāk   ət-tension*
       DEM   DEF.SG-tension.F.SG
       'This tension'

(2)   *wāḥəd   la pression*
       INDEF   DEF.F.SG-pression.F.SG
       'One pression'

(3)   *xāyba   l'indifférence*
       horrible.F DEF.F.SG-indifference.F.SG
       'Indifference is horrible!'

4.1   Mixed constituents

Mixed constituents as in (1) are recurrent in Moroccan Arabic-French codeswitching. Spontaneous data give evidence of various cases of French noun insertion. As we can see in the following examples, French nouns can be modified by Arabic

system morphemes such as the determiners *wāḥəd əl-* (4), *əl-* and *hād əl-* (5) as well as by a zero article (6).

(4) *'ənd-ək*   *wāḥəd* *əd-durée*
    have-2SG INDEF DEF-period.F.SG
    'You have a certain period'

(5) *ma-'ṛəfna-ši*   *dāba* *l-principe*   *gāʕ* *à quoi sert*
    NEG-know-1PL-NEG now DEF-principle.M.SG never to what serve-3SG
    hād   *l-calcul*   *u* *hād* *t-txəṛbiq*
    DEM.M.SG DEF-calculus.M.SG and DEM DEF-nonsense.M.SG
    'We do not know the principle at all. What's the purpose of this calculus and this nonsense?'

(6) *association*   *bla*   *wṛāq* *ṛā-h*   *xaṭiṛa*
    association.F.SG without papers.PL COP-3F.SG dangerous.F.SG
    'An illegal association is dangerous! '

4.2   Internal EL islands

Internal EL islands are EL (i.e. French) constituents integrated into a ML (i.e. Arabic) DP. This DP can be headed by the Arabic demonstratives *dāk* and *hād*, as well as by the determiner *wāḥəd* as we can see in the following examples:

(7) *'ṛef-ti*   *ana* *xāl'a-ni*   *bḥāl* *hād* *la*   *situation*
    know-2F.SG 1SG scare.3M.SG-1SG like DEM DEF.F.SG situation.F.SG
    'ṛef-ti   *žāb*   *li-yya* *wāḥed* *la*   *loi*
    know-2F.SG bring.3M.SG to-1SG INDEF DEF.F.SG law.F.SG
    ḥeṭṭ-ha   *'li-yya*
    put.3M.SG-3F.SG on-1SG
    'You know, as for me, a situation like this scares me. You know, he came up with a law and applied it on me.'

(8) *kāyen*   *wāḥəd* *les sujets*
    EXS.3M.SG INDEF DEF.M.PL-subjects-M.PL
    'There are some topics.'

Although internal EL islands are frequent in the corpus, they are not varied as they conform to the French NPs headed by Arabic definite articles that have already been documented in the literature (Bentahila & Davies 1983, 1998; Heath 1989; Lahlou 1991; Sankoff & Nait M'barek 1988).

4.3   EL islands

EL islands are well-formed maximal projections. When occurring into a CP, they conform to the grammar of the ML arguments. The examples bellow show how a French NP can be embedded into a bilingual CP by means of system morphemes such as the definite article *le*, the quantifier *chaque* 'every' and the possessive *notre* 'our'.

(9)   *ḥāwl-at   t-ḥəll-na   le problème*
      try-3F.SG   3F.SG-solve-1PL   DEF.M.PL-problem.M.SG
      'She tried to solve the problem for us.'

(10) *chaque semestre   ʻənd-na   programme*
      every semester.M.SG   have-1PL   program.M.SG
      'Every semester, we have a specific program.'

(11) *notre   pouvoir   d'achat   ġādi yə-hbəṭ*
      POSS.1PL power.M.SG of purchase.M.SG FUT   3M.SG-decrease
      'Our purchasing power will decrease.'

EL islands may include an adjective modifying a noun. Alike in French, the adjectives may be placed before or after the noun:

(12) *un faux   geste   yə-qdər*
      INDEF.M.SG-wrong.M.SG gesture   3M.SG-can
      y-dīr   bəzzāf d-les problèmes
      3M.SG-do a_lot   POSS-DEF.PL-problems.M.PL
      'A wrong reaction could cause a lot of problems.'

(13) *bḥāl 'ila ʻənd-ək   un   appareil   thermique*
      as   if   have-2SG INDEF.M.SG device.M.SG thermal
      'As if you had a thermal device.'

5.   Bilingual DP: How specific is it in Moroccan Arabic-French contact?

The previous examples reveal the diversity of nominal insertion through three structures: mixed constituents, internal EL Islands and EL islands in Moroccan Arabic-French codeswitching. However, a comparison of these three structures shows that they do not present the same distribution.

First of all, nominal insertions in which complex determiners keep their original form (*hād əl-* and *wāḥəd -əl*) are much less frequent. In this context, *wāḥəd* and *hād* tend to introduce French NPs. Secondly, it is worth noting that French articles inserted in the nominal structure headed by *wāḥəd* or *hād* also present

different distributions, *waḥəḍ le* N and *hād le* N being significantly less recurrent than *wāḥəd / hād la* or *les* N. Most importantly, this kind of insertion is recurrent and productive also when French is the ML, as showed by the following examples.

(14) *je   sens   wāḥəd-la froideur   f-dāk   la personne*
      1SG feel   INDEF-DEF.F.SG-coolness.F.SG in-DEM.M.SG DEF.F.SG-person.F.SG
      'I feel some coldness in that person.'

(15) *J'avais   des problèmes   avec wāḥəd la fille   à Rabat*
      have.1SG INDEF.PL-problem.M.PL with INDEF DEF.F.SG girl.F.SG in Rabat
      'I had some problems with a girl in Rabat.'

(16) *Il m'a   choqué hād   le semestre*
      3M.SG 1SG have.3SG shock   DEM DEF.M.SG-semester.M.SG
      'He has shocked me this semester.'

(17) *Je suis contre   hād   l'agressivité*
      be.1SG against DEM DEF.F.SG-aggressiveness.F.SG
      'I am against this aggressiveness.'

The same phenomenon has been already documented in other studies of Moroccan Arabic-French codeswitching during the last thirty years.

(18) *elle t'envoie   wāḥəd le liquide*
      3F.SG 2SG send INDEF DEF.M.SG-liquid.M.SG
      'She sends you a liquid.'   (Bentahila & Davies 1983: p. 304)

(19) *Ils   n'ont pas   dīk   l'assistanat*
      3PL NEG-have.3PL-NEG DEM DEF.M.SG-assistantship.M.SG
      'They don't have this assistantship.'   (Barillot-Fadel 2001: p. 272.)

Interestingly enough, our data also present new kinds of determiners which have never been attested so far. This is the case of French nouns introduced by definite articles in combination with other morphemes such as possessives (*mon*, *sa*), indefinite articles (*une*), quantifiers (*huit*) and indefinite determiners (*quelques*). These French NPs also occur with the complex determiners *wāḥəd* and *hād*.

(20) *wāḥəd une   distance   kbira*
      INDEF INDEF.F.SG distance.F.SG long.F.SG
      'A long distance.'

(21) *wāḥəd huit représentants*
      INDEF eight representative.M.PL
      'Some eight representatives.'

(22) *wāḥəd mon frère ġādi y-ṣāfəṛ*
     INDEF POSS.M.SG brother FUT 3M.SG-travel
     'A brother of mine will travel.'

(23) *walakin fina huwwa hād mon genre*
     but where 3M.SG DEM POSS.M.SG kind.M.SG
     'But where is my kind (of men)?'

(24) *t-xəyyli wāḥəd sa cousine kan-ət m'ā-h avant*
     imagine.IMP.F INDEF POSS.F.SG-cousin.F.SG be-3F.SG with-3M.SG before
     'Imagine her cousin was dating him before.'

(25) *ġāda l-wāḥəd un mois*
     go to-INDEF INDEF.M.SG month.M.SG
     'That goes back to a month approximately.'

(26) *kān 'ənd-i wāḥəd quelques relations*
     be.3M.SG at-1SG INDEF INDEF.PL-relationship.F.PL
     'I had some relationships.'

In the light of these examples, a number of questions arise. First, why do French determiners tend to modify French nouns after *wāḥəd* or *hād*? Furthermore, why are internal EL islands much more frequent than mixed constituents when Moroccan Arabic is the ML? And why are they frequent even when the ML is French? In the following paragraphs, I will try to find ways to answer to these questions.

6. Why do French determiners appear after *wāḥəd* or *hād*?

A first possible explanation is that French determiners in combination with French nouns are not instances of codeswitching, as they rather represent complex loanwords resulting from borrowing. This would explain why these structures resist to morphosyntactic integration and appear as insertion. In this regard, Poplack et al. (2015: p. 184) argue that "*because the syntactic constructions into which these DET+N chunks enter do not exist in FR, we conclude that on these measures they are behaving like borrowings, even though they are toward sequences rather than the canonical one.*" The comparison between Moroccan Arabic in contact with French and Dutch has led some scholars (Boumans 1998; Caubet & Boumans 2000; Muysken 2000, 2008; Nortier 1990) to propose a number of criteria to explain why French nouns keep their determiners while Dutch nouns are inserted as bare forms, as showed by the following examples:

(27) *ka-n-dəwwəz   wāḥəd cursus*
      PROG-1SG-follow INDEF   course
      'I follow a course.'   Moroccan Arabic-Dutch (Boumans 1998: p. 187)

(28) *dāk   handelscorrespondent*
      DEM   commercial_correspondent
      'That commercial correspondent.'
                                                   Moroccan Arabic-Dutch (Nortier 1990: p. 147)

Muysken argues that "*a superficial similarity between French le and the Arabic article l*" (2008: p. 171) is a factor favoring this structure, so if "*French le/la resembles Arabic l; Dutch de/het does not*" (2000: p. 83). He further explains that "*French articles are somehow treated as prenominal clitics (as is the case with the Arabic determiners themselves)*". Thus, they are obligatory and according to Muysen (2008: p. 171), they also appear "*as a component of the French noun in many French-lexicon creoles*". The morphological tightness between articles and nouns in both French and Moroccan Arabic may explain the frequency of the French defined NPs as compared to Dutch nouns which occur either as mixed constituents or bare forms (Boumans 1995: p. 58).

Myers-Scotton (2000: p. 118), on her part, observes that "*data from other codeswitching corpora [...] in which French is the EL provide a good reason to argue against the 'tie that binds' explanation for why French determiners can appear in Arabic/French codeswitching.*" For supporting this claim, she provides examples of Wolof – French codeswitching in which French nouns are integrated as bare forms due to the fact that, different from Arabic, Wolof has a post-nominal definite article:

(29) *am carnet   bi   seet ko*
      take notebook DET look 3SG
      'Take the notebook, look at it.'

(30) *bësal bouton bu   rouge bi*
      press button POSS red DET
      'Press the red button.'
                                             Wolof-French (L. Swigart 1992, in Myers-Scotton 2002: p. 118)

As we can see in the previous examples, the French nouns (*carnet/bouton*) are followed by the Wolof system morphemes (*bi, bu*). Accordingly, Myers-scotton (2002: p. 119) concludes that *"embedded language determiners (French here) can appear if they show sufficient congruence with their ML counterparts at all three levels of the abstract grammatical structure."* In this regard, it should be reminded that According to the 4-M models, two basic types of morphemes can be identified: content morphemes and system morphemes which can be further subdivided in early, bridge, and outsider morphemes. Determiners are early system morphemes,

as they are "*accessed at the level of mental lexicon*" (Myers-Scotton 2002: p. 17) and they "*are called early because of their early saliency*" (2002: p. 18). Early system morphemes appear at the lemma level and are activated together with the noun they modify, constructing a DP. So, if this construction occurs, it is because of the congruence between French and Arabic determiners at three levels of abstract grammatical structure (i.e. lexical-conceptual structure, predicate-argument structure, and morphological realization patterns,[1] Myers-Scotton 2002). Early system morphemes (i.e. determiners) do not need to come categorically from the ML (Myers-Scotton & Jake 2015: p. 434), as they can be provided by both French (EL) and Moroccan Arabic (ML). When they derive from French, we are dealing with a "*compromise strategy*" according to the MLF. However, the fact that the Arabic and French share a "*sufficient*" congruence of determiners on the three grammatical levels does not explain why internal EL islands involving El morphemes are more frequent than are mixed constituents.

7. Determiner phrase in Moroccan Arabic and French: A possible comparison

A comparison between monolingual DP in the two languages allows us to understand how the nominal insertion works in a situation of language contact. Determiners are functional morphemes with different properties in French and Moroccan Arabic. On the one hand, French stands out from other languages for the fact that it very rarely allows nouns to occur without an article. On the other hand, Moroccan Arabic could admit a zero article as determiners are obligatory and required in argument position (Poplack et al. 2015). There are three classes of determiners in French: definite, indefinite and particles. Definite and indefinite articles present three forms (masculine singular, feminine singular, masculine/feminine plural) and they always occur in pre-nominal position (Gary-Prieur 2011). Besides, a few other morphemes, like indefinite determiners or quantifiers, can be used for determining a noun.

According to Caubet (1993), there are four degrees of determination in Moroccan Arabic. These are respectively expressed by zero article, the definite and invariable *əl-*, the indefinite *ši*, and the complex determiner *wāḥəd əl-*. The functions of these forms cannot be reduced to a simple opposition between definite

---

1. "*Lexical –conceptual structure has to do with speakers' intentions regarding the meanings that they wish to communicate. Predicate –argument structure refers to syntactic structure and how thematic relations between lexical elements are realized. Morphological realization patterns refer to elements on the surface level*" (Myers-Scotton & Jake 2015: p. 436).

and indefinite NPs (Caubet 1993: p. 185). At the morphosyntactic level, the basic DP structure in Moroccan Arabic and French can be resumed as follows: French, e.g. *la solution* "the solution", DP [la-NP [solution]]; Moroccan Arabic, e.g. *əl-ḥall* "the solution" DP [əl-NP [ḥəll]] (R. E. Post 2010: pp. 31–33). As a further matter, Moroccan Arabic allows the determiners *hād/wāḥəd* to combine with the definite article *əl-* in the following way: DP [hād DP [l-NP [ḥəll]]]; DP [wāḥəd DP [l-NP [ḥəll]]]. French, on its part, presents few complex determiners in which definite articles (*le*, *la*, *les*), demonstratives and possessives can combine with indefinite determiners like (*quelques*: some, *tout*: all …), as we can see in the following example.

(31) *ka-t-kūn   une sortie   fi-ha   toutes   les promos*
      PROG-3F.SG-be INDEF.F.SG-outlet in-3F.SG all.F.PL DEF.PL.-discounts.F.PL
      'There is an outlet with all discounts.'

Other combinations are not allowed. For example, indefinites, demonstratives, possessives can be never associated with definite articles as in Arabic (DEF * DEM) or (* INDEF DEF / DEF): (* *Cette la solution*).

In the next section, I will focus my attention on the DP features in Moroccan Arabic-French codeswitching. *wāḥəd əl-* and *hād əl-* will be discussed separately.

8.   DP in Moroccan Arabic-French codeswitching: The indefinite *wāḥəd əl-*

Determiners in the Romance languages such as French are inflected for gender and number. Articles in French have three features: [+/– definiteness], [+/– gender] and [+/– number]. In contrast, Moroccan Arabic article does not encode number and gender; therefore, it has the following features: [+/– definiteness], [–gender] and [–number]. This fundamental difference may explain the occurrence of French determiners in Moroccan Arabic – French codeswitching (Myers-Scotton 2002; Muysken 2000). In both Moroccan Arabic and French the article belongs to the conceptual structure of the noun. When a French noun is selected at the lemma level, the article is also activated. Since there is insufficient congruence between both languages regarding gender/number, the French noun, which is a content morpheme, obviously requires definiteness shared by the EL and the ML. However, given that gender and number inflection is absent in the ML, the French noun appears with its article. This mismatch in the conceptual structure facilitates the formation of a maximal projection in French. Still, the grammatical dominance of the ML is strong, as the whole DP is constructed with the indefinite *wāḥəd*. Two factors may explain the high frequency of such combinations. The indefinite marker *wāḥəd* lacks two features ([+gender] and [+number]) which are instead specified

by the French noun. This can justify the frequency of internal EL and, at the same time, to explain the scarcity of the French noun insertions modified by *wāḥəd əl-*.

A comparison of how the DP functions when other languages are in contact turns to be useful for the present discussion. For instance, in Moroccan Arabic-English codeswitching, the complex determiner *wāḥəd əl-* is relatively frequent and the English noun is completely integrated:

(32) *dad ʿṭā-ni   wāḥəḍ al-book   mazinu*
      dad give.3M.SG-1SG INDEF DEF-book lovely.M.SG
      'Dad gave me a lovely book.'

(33) *šuf-t   waḥad al-big house kbira   wa zina*
      see-1SG INDEF DEF-big house big.F.SG and nice.F.SG
      'I saw a big house, it was big and beautiful'.
                          (Moroccan Arabic-English, Benchiba 2007: p. 240)

Grammatical gender is not assigned overtly to English nouns. So, like Moroccan Arabic, gender is not expressed by determiners. Palestinian, Jordanian or Iraqi Arabic in contact with English presents the same characteristic. First, as MLs, these idioms do not have composite determiners in their determination system (Brustad 2000), and English lacks determiners gender agreement. The following examples show how English noun can appear with Arabic system morphemes.

(34) *il-communities   ṣ-ṣġayyara   tʿallam-u stiʿmāl l-manure*
      DEF-community.PL DEF-small.F.SG know-3PL using DEF-manure.SG
      'The small communities learned using the manure.'
                          (Jordanian Arabic-English, Mustafa & Al-Khatib 1994: p. 221)

(35) *al-pain   yi-zīd*
      DET-pain 3M.SG-increase
      'The pain increases.'   (Iraqi Arabic-English, Sallo 1994: p. 124)

English nouns can also appear with a zero article. Whenever the article "the" appears, it is for building an EL Island, as we can see in the following examples.

(36) *idan ḥake-na bi-normal condition ʿil-intrapleural   pressure qaddeš*
      then say-1PL in-normal condition DEF-intrapleural pressure how much
      'Then, how much did we say is the intrapleural pressure in the normal condition?'   (Jordanian Arabic-English, Mustafa & AL-Khatib 1994: p. 220)

(37) *tarak-it-ha   in the car*
      leave-3F.SG-3F.SG in the car
      'She left her in the car.'   (Arabic-English, Rouchdy 2013: p. 137)

(38)   *na-xudh al-ināt*   *in the early morning*
       1PL-take DEF-female.F.PL   in the early morning
       'We take the females in the early morning.'

                                        (Iraqi Arabic-English, Sallo 1994: p. 125)

Asymmetry between Arabic and English in the domain of determiners agreement is the main factor underlying nominal insertion as mixed constituents or EI islands. However, when we look at typologically similar languages like Spanish in contact with Moroccan Arabic (Vicente & Ziamari 2008: p. 464), Spanish definite articles seem to appear frequently.

(39)   *dīk   el-niño   ma-bġa-ši   ya-kūl*
       DEM   DEF.M.SG-child NEG-want.3M.SG-NEG 3M.SG-eat
       'That child does not want to eat.'

In the previous example, the Spanish NP (*el niño*) functions in the same manner as French NPs. However, Spanish, as ML in contact with English, imposes its gender and number constraints as in the following example provided by Myers-Scotton (2015: 439).

(40)   *viene   mi familia   para los   holidays*
       come.3SG POSS family.F.SG for   DEF.M.PL holiday.PL
       'My family is coming for the holidays'

Gender and number are operational categories in Moroccan Arabic-French codeswitching (Treffers-Daller 1994: p. 123). This factor can explain why French nouns inserted in mixed constituents are very rare in my data. In this connection, it should be stressed that when a noun is masculine in French, it has more chances to be integrated with the Moroccan Arabic *wāḥəd əl-*.

(41)   *wāḥəd l-psychiatre*
       INDEF   DEF-psychiatrist.M.SG
       'A psychiatrist'

(42)   *ka-yə-qləb   m'a-k   l-méthode*
       PROG-3M.SG-change with-3M.SG DEF-way.F.SG
       'He changes the way he is behaving with you.'

In addition to gender/number and definiteness, "*there is a need for caution in seeking to attribute such differences directly to formal features of the languages involved*" (Bentahila et al., 2013: p. 328). Therefore, it is important to go beyond formal characteristics to explain these structures. Indeed, the pragmatic and enunciative dimensions are also necessary (Ziamari 2009).

It has been repeatedly observed that *wāḥəd əl-* fulfills particular discoursive functions in Moroccan Arabic (Caubet 1993, Brustad 2000). Caubet & Boumans (2000: p. 152) overtly state that "*the terms 'definite' and 'indefinite' are merely loose characterizations which do not cover all functions of these articles*". Furthermore, Caubet (1993: p. 267) observes that *wāḥəd əl-* fulfills two functions. On the one hand, it introduces cardinality. On the other hand, it expresses a qualitative determination. This article operates on discontinue category and conveys the operation of extraction of an element from a class. This semantic value is also expressed in French either by means of an indefinite or by a partitive determiner.

(43) *'ǝnd-kūm wāḥəd le gène   le sexe   féminin*
        at-2PL   INDEF DEF.M.SG-gene.M.SG DEF.M.SG sex.M.SG feminine.M.SG
        žbətt dāba le sabre   dyāl-i   'āyən n-gūl   l-kūm
        draw DEF.M.SG-sword POSS-1SG wait 1SG-tell to-2PL
        'ǝnd-kūm wāḥəd le gène   f-les   chromosomes
        at-2PL   INDEF DEF.M.SG-gene.M.SG in-DEF.PL chromosome.M.PL
        dyāl-kūm dima   xāṣ-kum   t-kūn-u dominées
        at-2PL   always must-2PL 2-be-PL dominated
        'You have a gene you females. I drew my sword, wait till I explain to you. You have a gene in you chromosomes; you always have to be dominated.'

In the previous example, *wāḥəd əl-* is used to extract one element from a class of genes. This value requires the indefinite article (*un*) in French. This asymmetry justifies the use of *wāḥəd əl-*. Undoubtedly, the semantic value of *wāḥəd əl-* does not invalidate the structural one, but it helps to understand why this structure is dominant in Moroccan Arabic-French codeswitching. Moreover, the indefinite marker provides a modal value. Caubet (1993 II: p. 287) argues that *wāḥəd əl-* expresses an "approximate assessment" ("*valeur d'approximation*") nuancing its quantitative degree. In such a context, the speaker is not sure about quantity and seeks to stress an approximate value. The examples bellow illustrate such modal interpretation.

(44) *ʕṛəf-ti   t-tamāra   lli dəwwəz-t   dāk l-ʕām*
        know-2F.SG DEF-ordeal.F REL spend-1SG DEM DET-year
        xəlli-ha   ʕla l-ḷḷāh t-tamāra   l-psychologique
        let.IMP-3F.SG on god   DEF ordeal.F DEF-psychological
        wāḥəd la pression   je me dis certainement je vais garder
        INDEF DEF.F.SG pressure.F.SG 1SG tell certainly   1SG FUT keep
        les séquelles   de tout ce que j'ai enduré hna f-l'ENSAM
        the aftereffects of all   1SG endured   here in-DEF-ENSAM
        c'est incroyable wāḥəd la pression   et hād l-ɛām
        unbelievable   INDEF DEF.F.SG pressure.F.SG and DEM DEF-year

> *je sens que je suis en train de revivre la même chose*
> I feel   that I am reliving   the same thing
> 'You can't imagine the ordeal I was through last year, forget it. It was a psychological ordeal. A pressure! I tell myself certainly I am going to keep the aftereffects of all that I endured here in the ENSAM. Unbelievable a certain pressure and this year I feel that I am reliving the same thing.'

The French determiners "*le/la*" or "*un, une*" cannot convey such semantic nuances. Therefore, semantic incompatibility between French and Moroccan Arabic can explains the recurrence of *wāḥəd* introducing a defined French NP.

9.   Demonstratives

In Moroccan Arabic, the definite article *əl-* is typically coupled with a demonstrative determiner like *hād*, *dāk*, *hādāk*. Thus, it is not surprising that in Moroccan Arabic-French codeswitching, Arabic demonstratives appear frequently in mixed constituents or internal EL Islands regardless of the ML. In such a context, French demonstratives are quite rare in EL islands. In my data, only one example of French demonstrative is attested in 11 hours of conversation:

(45)   *bien sûr n-šuf   un prof   ma-ġādi-š nə-ḍḥək de cette façon*
       of course 1SG-see INDEF.M.SG professor NEG-go-NEG 1SG-laugh in this way
       'Of course, if I see a professor, I won't laugh like this.'

According to Myers-Scotton (2002: p. 123), demonstratives behave like complex determiners. In addition, demonstratives are early system morphemes that can be provided by both languages. The present study proposes an additional argument seeking to go beyond morphosyntactic analysis.

The demonstrative determiner fulfills two functions: it plays a determination role (extraction) and a deictic function or pinpointing operation (Caubet 1993 II: p. 297). In Moroccan Arabic, a demonstrative always modifies a defined noun when it functions as a determiner and this seems to play a role in the occurrence of mixed nominal constituents and internal EL islands. Two scenarios are possible: either mixed constituents are built with a French noun satisfying the congruence test or an internal EL island occurs. In the latter case, we have to take into consideration to the main semantic features of determiners, namely definiteness, gender, and number. The study of the distribution of demonstratives shows that asymmetry between Moroccan Arabic and French is responsible for various structures in the corpus. The combination of some determinants by both languages, as it has been argued by Myers-Scotton (2002), may explain such results. The difference between

demonstratives in French and Moroccan Arabic on the semantic and pragmatic levels helps to bring new arguments in order to efficiently analyze this striking phenomena in codeswitching. However, if the concept of congruence, as conceived by Myers-Scotton, manages to explain a number of complex NPs, it is far from being universally valid.

10. Conclusion

This study examined the occurrence of internal EL islands in relation to nominal insertions in Moroccan Arabic-French codeswitching. This particular structure involves bilingual DPs in which French nouns tend to keep their definite article regardless of the ML governing the construction. Adopting the MLF model, the study showed that definiteness, gender and number are important features that can explain the occurrence of defined French NPs headed by the Moroccan Arabic indefinite marker *wāḥəd*. Besides, it has also been argued that demonstratives may fulfill the same inflectional functions, in combination with their original deictic and modal values. Above and beyond, the study brings to the fore the necessity of analyzing asymmetry in nominal determination, not only in the light of formal structures of the languages in contact, but also in consideration of a number of semantic and discursive factors.

List of abbreviations

COP	copula	INDF	indefinite
DEF	definite	M	masculine
DEM	demonstrative	NEG	negation
EXS	existential	PL	plural
F	feminine	POSS	posessive
FUT	future	PROG	progressive
IMP	imperative	SG	singular

References

Aabi, M. (1999). *The Syntax of Moroccan Arabic/French and Moroccan Arabic/standard Arabic code switching*. Unpublished PhD thesis. University of Sheffield.

Barillot-Fadil, N. (2001). Code-switching arabe marocain/français: Remarques générales et aspects prosodiques. In C. Canut and Dominique Caubet (eds.), *Comment les langues se mélangent: codeswitching en Francophonie*. Paris: L'Harmattan. 171–192.

Benchiba, N. (2007). Language & Grammatical Gender in Moroccan Arabic and English codeswitching: A levelling phenomenon. In José Santaemilia, Patricia Bou, Sergio Maruenda, Gora Zaragoza (eds.), *International Perspectives on Gender and Language*. Valencia: Valencia University. 233–248.

Bentahila, A. and E. Davies (1983). The syntax of Arabic-French code-switching. *Lingua* 59: 331–343. https://doi.org/10.1016/0024-3841(83)90007-4

Bentahila, A. and E. Davies (1998). Codeswitching: an equal partnership?. In R. Jacobson (ed.), *Codeswitching worldwide*. Berlin-New York: Mouton de Gruyter. 25–49. https://doi.org/10.1515/9783110812190.25

Bentahila, A. and E. Davies (2007). Review of Myers-Scotton: contact linguistics: Bilingual encounters and grammatical outcomes. *Language in Society* 36 (3): 455–459.

Bentahila, A. and E. Davies and J. Owens (2013). Codeswitching and Related Issues. In J. Owens (ed.), *The Oxford Handbook of Arabic Linguistics*. Oxford University Press. 326–348.

Boumans, L. (1995). Le Rôle des morphèmes de fonction dans le code-switching. In P. Bos (ed.), *Langues du Maroc: Aspects linguistiques dans un contexte minoritaire*. Tilburg: Tilburg University Press. 51–88.

Boumans, L. (1998). *The syntax of codeswitching: analyzing Moroccan Arabic/Dutch conversations* ("Studies in Multilingualism 12"). Tilburg: Tilburg University Press.

Boumans, L. and D. Caubet (2000). Modelling intrasentential codeswitching: a comparative study of Algerian/French in Algeria and Moroccan/Dutch in the Netherlands. In J. Owens (ed.), *Arabic as a minority language*. Berlin: Mouton de Gruyter. 113–180. https://doi.org/10.1515/9783110805451.113

Brustad, K. (2000). *The Syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian and Kuwaiti dialects*. Washington, D.C.: Georgetown University Press.

Caubet, D. (1998). Alternance de codes au Maghreb: pourquoi le français est-il arabisé? *Alternance des langues et Apprentissage en Contextes Plurilingues, Plurilinguismes* 14, Paris: René-Descartes. 121–142.

Caubet, D. (1993). *L'arabe marocain*. Volume I & II. Paris-Louvain: Peeters.

Chan, B. (2009). Code-switching between typologically distinct languages. In B. E. Bullock and A. J. Toribio (eds.), *The Cambridge Handbook of linguistic Code switching*. Cambridge: Cambridge University Press. 182–198. https://doi.org/10.1017/CBO9780511576331.012

Gardner-Chloros, P. (2009). *Code-switching*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511609787

Gary-Prieur, M.-N. (2011). *Les déterminants du français*. Paris: Ophrys.

Heath, J. (1989). *From code-switching to borrowing. A case study of Moroccan Arabic*. London and New York: Kegan Paul International Ltd.

Lahlou, M. (1991). *A morpho-syntactic study of code switching between Moroccan Arabic and French*. PhD thesis, Austin: The University of Texas.

Mustafa, Z. and M. Al-Khatib (1994). Code-mixing of Arabic and English in teaching Science. *World Englishes* 13 (2): 215–224. https://doi.org/10.1111/j.1467-971X.1994.tb00308.x

Muysken, P. (2000). Bilingual speech. *A typology of code-mixing*, Cambridge: Cambridge University Press.

Muysken, P. (2008). *Functional categories*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511755026

Myers-Scotton, C. (1993). *Duelling languages: grammatical structure in codeswitching*. Oxford: Clarendon Press, (1997). Afterward-paperback edition.

Myers-Scotton, C. (2002). *Contact linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
https://doi.org/10.1093/acprof:oso/9780198299530.001.0001

Myers-Scotton, C. and J. Janice (2015). Cross-language asymmetries in code-switching patterns Implications for bilingual language production. In J. W. Schwieter (ed.), *The Cambridge Handbook of Bilingual Processing*. Cambridge: Cambridge University Press. 416–458.
https://doi.org/10.1017/CBO9781107447257.019

Nortier, J. (1990). *Dutch and Moroccan Arabic in contact: code switching among Moroccans in the Netherlands*. The University of Amsterdam.

Poplack, S., L. Sayahi, M. Nahed and N. Dion (2015). An Exception to the Rule? Lone French Nouns in Tunisian Arabic. *Working Papers in Linguistics* 21 (2). University of Pennsylvania. 177–168.

Post, R. E. (2010). *Code-switching in the Determiner Phrase: A comparison of Tunisian Arabic-French and Moroccan Arabic-French switching*. Master of Arts, University of Texas.

Redouane, R. (2005). Linguistic Constraints on Codeswitching and Codemixing of Bilingual Moroccan Arabic-French Speakers in Canada. In C. James, K. T. McAlister, K. Rolstad and J. MacSwan (eds.), *Proceedings of the 4th International Symposium on Bilingualism*. Somerville: Cascadilla Press. 1921–1933.

Rouchdy, A. (2013). Language conflict and identity: Arabic in the American diaspora. In Alia R (ed.), *Language contact and language conflict in Arabic: variations on a sociolinguistic theme*. Londres-New York: Taylor & Francis. 133–148.

Sallo, I. K. (1994). Code-switching at the university: A sociolinguistic study. In R. de Beaugrande, A. Shunnaq and M. H. Heliel (eds.), *Language, discourse and translation in the West and Middle East*. Amsterdam: Benjamins. 115–132. https://doi.org/10.1075/btl.7.18sal

Sankoff, D. and M. Nait M'barak (1988). Le discours mixte arabe/français: emprunts ou alternances de langue? *Revue Canadienne de Linguistique* 33: 143–154.

Swigart, L. (1992). *Practice and Perception: Langauge use and attitudes in Dakar*. Ph.D. Seattle: Unversity of Washington.

Treffers-Daller, J. (1994). *Mixing Two Languages: French-Dutch Contact in a Comparative Perspective*. Berlin New York: Walter de Gruyter. https://doi.org/10.1515/9783110882230

Vicente Á. and K. Ziamari (2008). L'arabe marocain au contact du français et de l'espagnol. S. Procházka and V. Ritt-Benmimoun (eds.), *Between the Atlantic and Indian Oceans: studies in contemporary Arabic Dialects*. Vienne: Neue Beiheftezur Wiener Zeitschriftfür die Kunde des Morgenlandes Bd. 4. 457–469.

Ziamari, K. (2007). Development and linguistic change in Moroccan Arabic French codeswitching. In C. Miller, E. Al Wer, D. Caubet and J. Watson (eds.), *Arabic in the City*. London-New York: Routledge. 275–290.

Ziamari, Karima (2008). *Le code switching arabe marocain-français au Maroc: analyse linguistique*. Paris: L'harmattan.

Ziamari, K. (2009). Moroccan Arabic-French codeswitching and information structure. In J. Owens and A. Elgibali (eds.), *Information Structure in Spoken Arabic*. London-New York: Routledge. 189–206.

Ziamari, K. (2012), Le code switching arabe marocain/français: Quelle place pour une seule langue matrice? In A. Barontini, C. Peirera, A. Vicente and K. Ziamari (eds.), *Dynamiques langagières en Arabophonies: variations, contacts, migrations et créations artistiques, Hommage offert à Dominique Caubet par ses élèves et collègues*. Estudios de Dialectología Árabe, Saragosse: University of saragosse. 347–365.

Ziamari, K. (2013). Arabe marocain et français en contact: description morphosyntaxique et comparaison typologique. *Travaux du Claix*. Université de Provence. 337–348.

Ziamari, K. and A. Backus (2006). Embedded language Islands in Moroccan-Arabic French and Turkish-Dutch Codeswitching. *International Symposium on Cities of the Mediterranean: Old World Modernities in the New Millennium.* Cadiz.

From Arabia to Persia and back

Code-switching among the Āl ʿAlī tribe in the UAE and Iran

Dénes Gazsi
The University of Iowa

This paper explores the discourse functions of Arabic-Persian code-switching and the phonological/lexical outcomes of language contact among members of the Āl ʿAlī tribe in the United Arab Emirates and Hurmuzgān Province in Iran. The linguistic environment among the Āl ʿAlī is characterized by bilingualism and multidialectalism. In the spoken and written code, they generate a tetra-glossic switching between Modern Standard Arabic, Gulf Colloquial Arabic, Modern Standard Persian, Colloquial Persian and two Persian dialects: Bandarī and Ačumī. The study draws on recorded data with tribal members in the UAE and conversation threads of fellow Iranian tribesmen on social media sites. The main theoretical construct applied for the analysis is the Matrix Language-Frame model (Myers-Scotton 2002). It will be argued that the nature of codeswitching among the Āl ʿAlī is situational and transactional, both inter- and intra-sentential. Language and dialect choice is determined by the topic of the conversation, the interlocutors' identity and their relationship to each other.

Keywords: code-switching, bilingualism, multidialectalism, Persian Gulf, Gulf Arabic, Arabs in Iran, Matrix Language-Frame model, Persian dialects, Āl ʿAlī tribe

1. Introduction

The traditions of fishing, pearling, shipbuilding and long-distance trade on the waterways of the Persian Gulf have for centuries served as a vehicle for cultural exchange among the diverse ethnic and linguistic communities in the region (Nūrbakhš 2003: p. 10). From the numerous Arab tribes who still inhabit coastal villages in Iran, the Āl ʿAlī is representative of the Arabic-Persian bilingualism which, combined with diglossia, creates a dynamic situation of language contact between the two coasts. This paper is an attempt to examine the discourse functions of

code-switching, and to describe the phonological and lexical outcomes of language contact among members of the tribe in the United Arab Emirates and southwest Hurmuzgān Province on the Iranian shore of the Persian Gulf.[1] Using a wide variety of data, the study investigates the language profile and dual identity of Āl ʿAlī tribesmen by analyzing multiple cases of oral and written code-switching. No language data has thus far been recorded specifically for this tribe, and virtually none exists for other Arab tribes on the Iranian littoral. Though the current paper is restricted to the Āl ʿAlī, and therefore does not purport to be a comprehensive linguistic analysis of Iranian coastal Arabs, it will highlight how they generally project their ethnic minority status through the use of Arabic and Persian language. From a participant-centered perspective, it will also compare the sociolinguistic aspects of their bilingualism with those Āl ʿAlī members who already relocated to the Arabian Peninsula. The main theoretical construct I apply for the analysis is the Matrix Language-Frame model (Myers-Scotton 2002). The MLF model is the dominant model of insertional code-switching and the most developed model for explaining bilingual constituents. The model posits that there is asymmetry between the participating languages with regard to their roles. In classic code-switching, only one language, the Matrix Language (ML), supplies the morphosyntactic frame of the clause (these are system morphemes). The other participating language, the Embedded Language (EL), supplies content morphemes that assign or receive thematic roles.

After a brief outline of the terminology and historical background of Iranian coastal Arabs, I discuss the extent of the Āl ʿAlī's rulership. This is followed by a description of their language use before I move on to the language data. I will investigate how the MLF model conforms to their language situation, particularly to patterns of code-switching. The central methodology in this paper is code-switching in the spoken as well as the written language, an emerging area of study that seeks to complement the spoken data in an era when electronic media has significantly changed the way we interact with each other. As people of all ages increasingly spend an equal amount of time communicating in the virtual and the real world, using oral and written materials increase the range of data sets, and add a new horizon to sociolinguistic and dialectal analyses.[2]

---

1.   Together with Būšihr Province further north, Hurmuzgān extends along the southeastern section of the coastline in Iran, with Bandar ʿAbbās as capital and largest city.

2.   In recent years, social media has steadily begun to blur the boundaries between a language's spoken and written registers. The pervasive appearance of Arabic and Persian dialects in chatrooms and online conversation threads presents a solid shift towards their tolerability as a code of public communication (Sayahi 2014: p. 79).

2.　Arabs on the Iranian Gulf Coast

Few in-depth studies have been published on the past and contemporary Arab presence on the Iranian coast of the Persian Gulf. The topic's absence from the academic realm can be ascribed to several reasons, primarily to the political and cultural sensitivity of the issue, but also to the obscurity this group has fallen into since the mid-20th century CE.[3] Despite the scarce information that exists about the current status of this population, the migration of Arab tribes from Eastern Arabia during the 18th and 19th centuries CE, and their establishment of villages on the opposite coast were extensively documented in British government records. Iranian anthropological, cultural and geographical literature mentions Arabs from a historical perspective, but their comments on contemporary life of Arabs seldom go beyond the observation that the long white Arab robe for men and the traditional face mask for women (*burqaʿ*) are habitually seen in Hurmuzgān Province (Bakhtiārī 2001: p. 56). Likewise, in Western academic publications, Iranian coastal Arabs tend to be regarded as a relic of the past (Nadjmabadi 2009: p. 139, Floor 2014: p. xiv). The word *Hōla* (variously referred to as *Hula*, *Hawala*, *Huwala*) is the historically recognized name for Arabs on the Iranian coast of the Persian Gulf and those returned to Arabia (Al-Dailami 2014: p. 301, Floor 2014: p. 19). However, contemporary Arab inhabitants of Iran's coastline do not accept this term as legitimate. They view it as one fostering the assumption that they are Arabized Persians, and not descendants of Arabian tribes.[4] Instead, their preferred endonyms include ʿ*Arab al-Juzur wa al-Sāḥil al-Šarqī li al-Khalīj al-ʿArabī* 'Arabs of the Islands and the East Coast of the Arabian Gulf', ʿ*Arab al-Sāḥil al-Šarqī* 'Arabs of the East Coast', ʿ*Arab al-Sāḥil* 'Arabs of the Coast' or ʿ*Arab Barr Fāris* 'Arabs of the Persian Land'. The undiscovered history of the coastline in the second part of the 20th century provides a good reason why social media can be a starting point for acquaintance with Iranian Gulf Arabs in general, and the Āl ʿAlī in particular. The internet serves as a networking forum for the two communities, those remaining in Iran and their relocated fellow tribesmen in Arab Gulf States. Many young Āl ʿAlī members as well as coastal Arabs from other tribes are making a concerted effort to maintain or reclaim their Arab roots and identity by engaging in a surging online activity. Since 2012, I have been following the social media presence of the Āl ʿAlī, and have monitored the progress of their auto-documentation and outreach. This is part of a comprehensive exploratory work to uncover the cultural, ethnic and linguistic

---

3.　This is in contrast to the recognition of Arabs in Iran's Khūzistān Province and the availability of academic work on their history, culture and language (Gazsi 2008: p. 195).

4.　See the distinction between 'old' Hawala and 'new' Hawala in Al-Anṣārī (2014: p. 375).

heritage of Arabs on the Iranian coast of the Persian Gulf, and chronicle their contacts with the Arabian Peninsula. Simultaneously, I recorded language data from both the resettled and visiting Āl ʿAlī members in the United Arab Emirates in 2011, 2013, 2015 and 2016. Speakers included ordinary and senior tribal members born in Iran but residing in Abu Dhabi, Sharjah and Dubai, and fishers and merchants who entered Emirati waters from the Iranian Āl ʿAlī territories on motor boats to unload their catch at the Fish Market in Ajman.

3.   The Āl ʿAlī and their dominion

The Āl ʿAlī, or less frequently al-ʿAlī, is a tribal federation spread across parts of the Arabian Peninsula and the Shībkūh region of the Islamic Republic of Iran.[5] The tribe's origins can presumptively be traced back to the Muṭair confederacy of Najd, though the exact lineage is mostly legendary and questionable (Al-Anṣārī 2014: p. 307). During the 18th century CE, Khalfān b. ʿAmrū moved to the Gulf coast where his son, Mājid, established the al-Muʿallā branch of the tribe, the current rulers of the Emirate of Umm Al-Quwain. His brother, ʿAlī, crossed the waterways to Persia with several families of the Āl ʿAlī, settled down in villages in the Shībkūh area, and extended his influence over the territory around Bandar Chārak and Kīsh Island (Floor 2014: p. 22). His successors became local sheikhs, and either fully or partially controlled this subregion until 1976. The entire ruling family was forced to leave Iran during the early days of the Islamic Revolution in 1979, and were repatriated into the Arab Gulf States. Their kin, the al-Muʿallā of Umm Al-Quwain expressed hostility toward them as they feared the newly arrived elders may want an equal distribution of wealth. Despite the intra-tribal animosity, Sheikh Zāyid b. Sulṭān Āl Nahyān, the founder of the UAE, granted Emirati citizenship to the returning members of the Āl ʿAlī. Ordinary people from the tribe, however, have remained in Iran's coastal villages and Kīsh Island where they live to this day. Nowadays, scanty economic resources on the coast are driving the Āl ʿAlī's younger generation to relocate to the United Arab Emirates, without much hope of acquiring citizenship. They retain their Iranian nationality, or upon expiration of their documents, they either join the growing number of stateless people (the

---

5.   Shībkūh, 'sloping mountain' in Persian, is a mountain range extending from Bandar Lengeh (Linja in Arabic) in the east to Pārsiān (formerly Gāvbandī) in the west, and from Kūkhird to Aškanān in the north. Prior to the mid-1900s, it was divided into four subregions, each ruled by an Arab tribe: the ʿUbaidalī tribe in the eastern, the Banī Bišr in the southern, the Āl Ḥammādī in the western, and the Āl ʿAlī in the northern areas, in addition to the Āl Ḥaram, al-Marāzīq and Āl Naṣūr (Lorimer 1970: II A, 62, Lorimer 1970: II B, 1782–3, Floor 2014: p. xii).

Bidūn), or purchase citizenship from the Comoro Islands. The Āl ʿAlī, being Arabs and Sunni Muslims of the Ḥanbalī school, constitute a minority in Iran from both an ethnic and religious perspective.[6]

Traditionally, the territory of the Āl ʿAlī on the Iranian coast has encompassed Bandar Chārak 'Chārak Port' with its surrounding villages and Kīsh Island, across the Gulf from Abu Dhabi (Floor 2014: p. 50). In 2011, Chārak had a population of 3,758 people in 711 households (*Saršumārī* 2011). Most residents are of Arab descent and bilingual speakers of Arabic and Persian. Ethnic Persians from other regions of Iran, most notable the Lurs, have only recently begun to settle into the town. Captain George Barnes Brucks of the East India Company's Bombay Marine, and author of one of the earliest known surveys of the Gulf, wrote between 1820 and 1830 that there were 900 men from the Āl ʿAlī tribe in Charrak (sic!), 360 of them fighters (Al-Anṣārī 2014: p. 207).[7] The remaining residents were fishermen and merchants. Today, Chārak is a sleepy fishing village, a shadow of the once glorious place that benefited from lying near one of the best pearling spots on the Iranian coast. Even the sheikhs' former fort, the only historic building in town, is in severe decay (Al-Anṣārī 2014: p. 325). Chārak is the primary access point to Kīsh Island from mainland Iran, and locals are employed in the passenger ferry business to and from the island (Nūrbakhš 2003: p. 77).

Kīsh, known for its lush palm gardens, is a historically prominent island with considerable influence since the Middle Ages. The island's Arabic name Qais, pronounced /gēs/ by the local Arabic-speaking population, may come from the name Jazīrat al-Qays b. ʿUmāra, first mentioned by Yāqut.[8] The island had a turbulent history, but since it is located only 20 km off the mainland, it engaged in a flourishing commercial and maritime activity between India, Persia and Iraq (Lorimer 1970: II B, 1471). Ḥasan b. Muḥammad b. Ṣāliḥ, the last Āl ʿAlī sheikh, relinquished his rulership in 1976, and the island gradually turned into a burgeoning tourist destination for Iranians who wanted to enjoy the relatively free atmosphere compared to the rest of Iran. Since Kīsh's designation as a free trade zone, luxury hotels and high-end shopping malls have cropped up at an incredible pace. The island's

---

6. This contrasts them with the Khūzistāni Arab tribes who are predominantly Shiite (Gazsi 2008: p. 196).

7. A detailed description of the village in the early 1900s is provided by Lorimer (1970: II A, 354).

8. Potts 2004. Potts' article provides a detailed history of Kīsh until the 14th century CE, then abruptly jumps to the Qajar period. He glosses over the first half of the 20th century CE, and concludes that the Kish Development Organization was founded in 1972. The entry makes no mention of the Arab connections of the island nor the immigration of the Āl ʿAlī during the Qajar era.

population in 2011 was 24,819 in 7,954 households, mostly Persians who moved to the island to work in the tourism industry. Thus, very few Arabs of the Āl ʿAlī tribe remain on the island, and their only neighborhood is Ṣafīn (also called Mīr Muhannā), a tiny enclave in the island's northwestern corner with traditional stone and mud houses.[9] Local Arabs operate a bazaar, the Bāzār-i ʿArabhā 'Bazaar of the Arabs', and a heritage house named Bait ʿAbdullāh bin Shāhīn.

4.   Language use among the Āl ʿAlī

The Arabic and Persian language are primary examples of diglossia. The high (H) and low (L) variety for Arabic are Modern Standard Arabic (MSA) and Colloquial Arabic respectively, in our case Gulf Colloquial Arabic (GCA). For Persian, they are Modern Standard Persian (MSP) and Colloquial Persian (CP). CP is the standardized form of the dialect of Tehran used throughout Iran in informal oral and written communication. Vernaculars divergent from CP abound in every region of the country. The sociolinguistic context on both shores of the Gulf is multifaceted, and the peculiarity of the language use of the Āl ʿAlī as an ethno-linguistic community lies in its perplexing complexity. Through continued interaction and occasional intermarriage with the Persian-speaking population on the Iranian littoral, the tribe represents a convergence of languages and dialects. MSP, the autochthonous and official language of the country is taught in schools and used in mass media, but commercial ties among settlements have prompted the Āl ʿAlī to communicate with Persians in the two main regional dialects, Bandarī and Ačumī. Bandarī is the coastal dialect group in Iran spoken by the native population of the port cities including Bandar Lengeh to Bandar ʿAbbās. Ačumī (*ačum* 'I go') is an umbrella term covering the vernaculars spoken in the coastal hinterland and mountainous regions of Southern Fārs and Northern Hurmuzgān Provinces with major centers in Bastak, Lār, Girāsh and Iwaz. Both groups belong to the South-Iranian dialectal area, and share many phonological, morphological and lexical features (Schmitt 1989: p. 295). Dialectal overlapping is common in the coastal speech communities, as many Ačumī speakers from the mountains have moved to live and work in the ports. They would speak their subdialect of Ačumī such as Lārī at home within the family, communicate with local people in Bandarī, and use MSP or CP with people outside their ethno-societal group (Pelevin 2010: p. 58).

---

9.   According to Bakhtiārī's geographical and anthropological description of Hurmuzgān Province, the inhabitants of Kīsh speak Arabic and Persian, in this order (Bakhtiārī 2001: pp. 148–9).

Persian multidialectalism in the coastal region is complemented with Arabic-Persian bilingualism among the Āl ʿAlī. Regardless of which side of the Gulf they live on, many members of the tribe are fluent speakers of MSP, know Bandarī and Ačumī to a varying degree, but most consider Arabic as their first language. In communication events, whether it be the spoken or written code, the Āl ʿAlī engage in bilingual and diglossic code-switching. They freely juxtapose the H and L varieties of Arabic and Persian in the same conversation or sentence. Despite the unavailability of public education in Arabic, educational bilingualism exists to a certain extent among the Āl ʿAlī on the Iranian coast. Classical Arabic morphology and syntax is taught throughout Iran in primary and secondary schools for the purpose of allowing students to read the Quran and other religious texts. Arabic remains a heritage language among the Āl ʿAlī on the Iranian coast, but their easy access to Arabic television channels and online newspapers has fostered Arabic-Persian bilingualism.

5.   Language data

The interviews I conducted with members of the Āl ʿAlī during field trips to Dubai, Abu Dhabi, Sharjah and Ajman span over a total of 25 hours. Altogether, 18 individuals agreed to provide spoken language data, both tribal leaders and ordinary tribal members in the age range of 19–51 years. All were born on the Iranian coast, but 9 relocated to the United Arab Emirates at various times of their lives. 3 speakers were visiting from Iran as fishermen or merchants, and 1 senior member of the tribe from Qatar was also present in Dubai. For the recorded data, I exclusively had access to men. The written data covering the years 2012–2015 is taken from the public social media accounts of 13 individuals of both genders (1 woman, 12 men) in the same age range. 2 people overlap in the two data sets as they consented to have their speech recorded in addition to allowing me to follow their computer-mediated communication with remaining fellow tribesmen on both shores of the Gulf. During the research, I abstained from interviewing members of the al-Muʿallā, the current ruling family of Umm Al-Quwain, as they decided to stay in Arabia and were hence not exposed to Persian influence. Due to the similarity of utterances produced in writing and oral communication, the following linguistic analysis combines the recorded data with the conversation threads and posts appearing on the participants' social media pages. A public group on social media created in October 2012 is dedicated to current and past Arab inhabitants of Chārak and serves as a gathering spot for auto-documentation and outreach. Its forty members regularly post pictures of local sites, famous people and upcoming events. They write extensive comments to these postings and engage in conversations regardless

of the distance between them. In addition to Chārak, members live on Kīsh, the UAE, and even in Shiraz or Tehran. Sentences and conversation threads from this group are highlighted with the acronym CHAF. I organize the language data under three headings: Persian monolingual data, Arabic monolingual data and bilingual data. As I move from one heading to the next, I will describe the language profile of three speakers in detail, one form each group outlined below, and demonstrate how MSA, GCA, MSP, Colloquial and dialectal Persian coexist and blend. Informants from the three groups are coded according to the following variables: residents of Chārak are marked with CH, Āl ʿAlī members from Kīsh with KI, and Iranian sheikhs-turned-Emirati citizens are listed as UA. The two-letter combination is followed by a number representing the individual speaker. To avoid confusion, I follow the rules of Arabic phonological transcription throughout the article. I frequently juxtapose Persian and Arabic sentences, but offering two sets of transcriptions would render the analysis incomprehensible. I will pinpoint instances when Persian words are pronounced with the Arabic phonemic inventory and vice versa. All texts are transcribed, but the online material remains in its original format and spelling, which occasionally includes typos.

5.1   Persian monolingual data

The Persian monolingual data exhibits features of spoken and written diglossic code-switching (between MSP and CP) and bidialectal code-switching (between CP and Ačumī), the latter being typical of Persian speakers across southwestern Iran. Moreover, the phonetic effect of Arabic sounds on the Persian speech of the Āl ʿAlī is widely attested in the recorded data. The scope of this phenomenon varies among speakers, and may affect one or two specific sounds in limited vocabulary, or may result in the full-fledged use of the Arabic vowel system with Persian words. While the Matrix Language-Frame Model as well as other models study code-switching on the morphosyntactic and lexical level, switching on the phonological level has been rarely analyzed systematically (Poplack et al. 1988).

(1)   CHAF: CH1 posted a question in CP, and the subsequent conversation occurred in MSP.

CH1: ایا حقیقت داره که نام قدیمی جارك جاه رقیه بوده

UA1: بله همه کتابهای تاریخ اینرا نوشتند اقای...

CH1: خیلي ممنون از شما بابت معلومات

UA1: خواهش میکنم عزیزم

CH1 [CP]:   *Āyā ḥaqīqat dāra ki nām-i qadīmī-yi Jārak Jāh Ruqayya būda.*

UA1 [MSP]:   *Bala hama kitābhā-yi tārīkhī īnrā niwištand āqā-yi …*

CH1 [MSP]:   *Khailī mamnūn az šumā bābat-i maʿlūmāt.*

UA1 [MSP]: *Khāhiš mī-kunam ʿazīzam.*
CH1: 'Is it true that the old name of Chārak was Chāh Ruqayya?'
UA1: 'Yes, all history books wrote this, Mr. [name].'
CH1: 'Thank you very much for this information.'
UA1: 'You're welcome, my friend.'

UA1 is the current leader of the Āl ʿAlī tribe, a distant cousin of Ḥasan b. Muḥammad b. Ṣāliḥ Āl ʿAlī, the last ruling sheikh of Chārak. He comes from the final generation of the Āl ʿAlī elite born on the Iranian shore. His father was an Arab, while his mother stemmed from an Ačumī Persian-speaking family in Bandar Lengeh. UA1, his six brothers and the remaining members of the ruling clan moved back to Arabia two years before the Islamic Revolution. Two of his brothers chose Qatar as primary residence, while he settled down in Dubai. Over time, he and his brothers seemed to have fully integrated into Gulf societies in terms of education, allegiance, dress and lifestyle, but through marriage UA1 still upholds family ties with Iran. UA1's language use is astonishingly complex. His Iranian homeland, upbringing and cultural environment has shaped his identity, and his active presence on social media continues to do so. He would speak MSP and CP with Iranian relatives, GCA with Emirati friends, and Urdu with Indian/Pakistani customers. Additionally, UA1 is fluent in both the Ačumī and Bandarī dialects of Persian. When talking to his brothers he prefers Ačumī, which he views as his true mother tongue, Arabic being only his 'father' tongue. UA1 is a proud Arab born on Iranian soil and proud Emirati citizen wearing traditional local dress, but one who speaks Ačumī Persian with his closest relatives.

CH1 apparently used an Arabic keyboard to type the word Chārak and *čāh* 'well, pit', where the Persian voiceless palato-alveolar affricate /č/ with three dots (چ) is substituted for the Arabic voiced palato-alveolar affricate /j/ (ج). The recorded data suggests that the Āl ʿAlī speakers clearly differentiate between /č/ ~ /j/ and /p/ ~ /b/, and can pronounce these sounds correctly.

In Examples (2) and (3), contact-induced language change is exhibited in the form of phonological borrowing, with cases of insertion of MSA sounds into CP.

(2) When UA1 discussed his family background in CP, he pronounced the words according to Persian phonological rules. Notable exceptions are: (1) the Arabic voiced pharyngeal fricative *ʿain* is pronounced in the word [ʕarab] 'Arab', not the equivalent Persian glottal stop [ʔarab]. Conversely, the word *ʿajamī* 'non-Arabic, Persian' is pronounced with the glottal stop [ʔad͡ʒami:]; (2) the first vowel in the word *aṣālat* conforms to the Arabic pronunciation, not the Persian with a *kasra* [esʋ:lat].
UA1 [CP]: *Mādar-am māl-i Linga hast. Mādar-iš ʿarab būd. Pidar-iš ham az ʿarabā yaʿnī aṣālat-išūn wakhtī nigāh mī-kunī az anṣāriā būdan az qadīm*

az Madīna. Anṣāriyā mī-gan az qadīm zamān-i paiġāmbar-i Islām muhājirat kardan ba Īrān. Walī ʿajamī ṣuḥbat mī-kardan.

UA1: 'My mother is from [Bandar] Lengeh. Her mother was an Arab. Her father was also from the Arabs, I mean, when you look at their origin, they were from the old Anṣār [= the Helpers], from Medina. It's said that the Anṣār long ago at the time of the Prophet immigrated to Iran. But they spoke ʿAjamī [= Persian].'

(3)   KI1 identifies himself as a pure Arab, and travels frequently between Kīsh and Ajman. He employs the Arabic vowel inventory in his CP speech: [kudu:m] 'which' for Persian [kodu:m]; [fa:rsi:] 'Persian' for Persian [fɒ:rsi:]. Though Persian nouns have no grammatical gender, KH1 resorts to a morphosyntactic borrowing by marking the Persian word *šahr* 'city' with the Arabic feminine ending to reflect its Arabic equivalent, *madīna*.

KI1 [CP]: *Kudūm šahra mī-šīnī?* 'Which city do you live in?'

[CP] *Fārsī khūndam walī fārsī dūst na-dāram.* 'I studied Persian but I don't like Persian.'

5.2   Arabic monolingual data

This section presents examples of the Arabic monolingual data occurring in both speech and writing. The Āl ʿAlī on the Iranian side of the Gulf speak a dialect similar to Emirati Arabic, while members who returned to the Arabian Peninsula either speak Emirati or Qatari Arabic, depending on where they chose to settle down. Despite the minor vernacular differences attributable to location, I opted for the more generic term Gulf Colloquial Arabic (GCA). Diglossic code-switching is widely known to occur between MSA and Arabic dialects, and Walters advocated that the Matrix Language-Frame model is applicable to the mixing between H and L (Walters 1996: p. 181). My data supports this claim.

(4)   CHAF: UA1 posted a question on social media where inter-sentential code-switching is attested. He uses MSA as the ML and the GCA expression *intu šū akhbārik* 'you all, how are you?' (lit. what are your news?) as the EL. The use of the 2nd pers. sg. masc. possessive suffix (-*ik*) after the 2nd pers. pl. personal independent pronoun (*intu*) may have been a typo. CH2's reply to the question is in GCA written with a crude Latin transcription.

UA1:   مرحبا شباب اليوم عندنا امطار غزيره من ساعة 7 الصباح و حتى الان و انتو شو اخبارك اليوم مع الامطار ؟؟؟؟؟؟؟؟؟

CH2: rohna ma andena shi pass zen ahsan men gabel shahrel jaE anshaalah Ekon ahsan

UA1 [MSA as the ML – GCA as the EL]: *Marḥaba šabāb al-yaum ʿindanā amṭār ġazīra min sāʿa 7 aṣ-ṣabāḥ wa ḥattā al-ʾān wa intu šū akhbārik il-yōm maʿa al-amṭār?????????*

CH2 [GCA]: rohna ma andena shi pass zen ahsan men gabel shahrel jaE anshaalah Ekon ahsan. [= *Rūḥnā mā ʿandinā šī bass zēn aḥsan min gabl šahr al-jāy inšallāh ikūn aḥsan.*]

UA1: 'Hello folks, today we had abundant rain from 7 in the morning until now, and you all, how are your news with the rain?'

CH2: 'We don't have any of that. But fine, [it's] better than earlier. Next month, God willing, it will be better.'

(5) In the speech of the tribe's ruling elite, contact-induced language change is manifest in the form of phonological borrowing from MSP/CP into GCA. These changes affect MSA interdentals: the voiceless interdental fricative *ṯ* [θ] is pronounced as a voiceless alveolar sibilant [s], the voiced interdental fricative *ḏ* [ð] is pronounced as a voiced alveolar sibilant [z], while the emphatic voiced pharyngealized alveolar fricative *ẓ* (zˤ) is also realized as [z]. These MSA sounds are retained in regular GCA, and since the phonological changes are not attested in either the Bandarī or Ačumī dialects of Persian, they are considered a direct MSP influence.

UA1: *fī ṣ-ṣaff as-sālis* 'in third grade'

UA2: *samān malāyīn* 'eight million'; *izā* 'if'; *atzakkar* 'I remember'; *bū zabī* 'Abu Dhabi', *abyaz* 'white'

(6) CHAF: This conversation demonstrates the anomalies of using GCA in written communication. CH4 posted a friend's photo who lives on Kīsh. The first comment comes from CH3, a man in his early 20s, and the driving force behind the online community of the Āl ʿAlī in Chārak. His profile states that he is a Sunni Muslim and knows 'Arabic, English and Persian'. His identity is clearly reflected in the order of languages, where the official language of his home country only comes in third.

CH3: مرحبا بی شباب گیس

CH4: گیس شو قیس

CH3 [GCA]: *Marḥaba bī šabāb Gais.*

CH4 [GCA]: *Gais šū Qais.*

CH3: 'Greetings, young people of *Gēs*.'

CH4: 'What *Gēs*? *Qēs*.'

CH3's comment sought to reflect the dialectal Arabic pronunciation of the island by typing it with the Persian /g/, where it is part of the phonemic inventory: گیس [*gēs*]. However, CH4 rebuffs this spelling, and corrects it to the MSA version with the /q/. If this conversation occurred in speech, no code-switching would be present.

However, the Iranian background of the speakers and the availability of Persian letters on the keyboard allowed the speakers to identify the dialectal realization of MSA /q/ as a separate phoneme in GCA. The awareness of phonetic differences between the two languages lead CH3 to superimpose the Persian spelling on the Arabic dialectal form.

5.3   Arabic-Persian bilingual data

Arabic and Persian have existed along a geographic continuum in the Gulf for over 1400 years. The uniqueness of language contact in the region is the overlap of diglossia and bilingualism. Bilingual code-switching in the written and oral communication of the Āl ʿAlī happens in different formats for various reasons. The choice of language is governed by the speakers' language preference, the nature of the topic, and the associations it induces. As most informants use both languages equally in monolingual contexts, switching between the two does not result in a systematic mixing. Both inter- and intra-sentential code-switching are attested, and the most conspicuous lexical categories borrowed from the other language or register are nouns and nominal phrases. In examples of intra-sentential code-switching, MSP frequently serves as the ML with CP, Bandarī dialect or MSA as the EL, but cases of (1) CP as the ML with Bandarī dialect as the EL, (2) MSA as the ML with GCA as the EL, and (3) GCA as the ML with Bandarī dialect as the EL also transpire. First, here are two instances of inter-sentential code-switching from social media, (7) and (8).

(7)   CHAF: A proverb in MSA was posted as a photo caption, followed by a comment in MSP by the same person. The language choice was influenced by the unavailability of an equivalent proverb in MSP.

CH4: ما يهزك ريح

CH4:   این یکی از لنجهای تندرو ایرانی هست به دلیل حمل سوخت زیاد پایین رفته

CH4 [MSA]: *Mā yahizzak rīḥ.*
CH4 [MSP]: *Īn yakī az lanjhā-yi tundrau-i īrānī hast ba dalīl-i ḥaml-i sūkht-i ziyād pāyīn rafta.*
CH4: 'The wind should not shake you.' (= May you sail safely.)
CH4: 'This is one of the fast-paced Iranian motor boats that drowned due to its heavy load.'

(8)   CHAF: MSP and MSA appear in immediate succession in the same written dialogue in the following brief exchange. The form in parenthesis in UA1's comment is the clarification of the spelling of Chārak with an Arabic /j/, as he was using an Arabic keyboard that did not include the Persian sound /č/ written with three dots (چ).

CH3: ساحل زیبای بندر چارک

UA1: یا سلام على ساحل جارك (تشارك)

CH3 [MSP]: *Sāḥil-i zībā-yi Bandar Chārak.*

UA1 [MSA]: *Yā salām ʿalā sāḥil Jārak (Tšārak).*

CH3: 'The beautiful coast of Bandar Chārak'

UA1: 'How nice is the coast of Chārak.'

(9) CHAF: UA1 regularly writes personal updates on his social media account in both Arabic and Persian to reach out to his friends on both coasts of the Gulf. Here is an example of a bilingual post followed by a conversation in monolingual sentences in GCA and MSP. Additionally, we see diglossic code-switching between MSP and Bandarī dialect, bilingual code-switching between MSP and MSA, and bilingual-diglossic code-switching between MSP, MSA and Bandarī dialect. The conversation thread elicits how a written dialogue endeavors to preserve the phonological, morphological and syntactic features of GCA.

UA1: رجعنا للبلاد فجر اليوم بعد رحلة استغرقت اسبوع == بس از یکهفته مسافرت سحرکاه امروز بر کشتیم بدبی

UA3: الف سلامه

UA1: الله ايسلمش و ايخليش يا …

CH5: رسیدن بخیر

UA1: ممنون خالو جان

CH6: خوش امدی علی عینی یا اخی کاک … طبق فرمان خدا سیروا فی الا رض را همیشه داشته باشی همراه با شادی و سلامت

UA1: قربانت کاکا …

CH7: انشالله که سفر خوبی توبوده بشه … و خوش آمد بشما عمي عزیز

UA1: خیلی ممنون عمی عزی خوب هسته جای شما خالی

UA1 [MSA == MSP]: *Rajaʿnā li al-bilād fajr al-yaum baʿd riḥla istaġraqat usbūʿ. == Bas az yak-hafta musāfirat saḥarkāh-i imrūz bar-kaštīm ba-Dubai.*[10]

UA3 [GCA]: *Alf salāma.*

UA1 [GCA]: *Allāh isallim-iš wa ikhallī-š yā …*

CH5 [MSP]: *Rasīdan ba-khair.*

UA1 [MSP]: *Mamnūn khālū jān.*

CH6 [MSP as the ML – MSA and Bandarī dialect as the ELs]: *Khuš āmadī ʿalā ʿainī yā akhī kāk … ṭibq-i farmān-i khudā sīrū fī al-arḍi rā hamīša dāšta bāšī hamrāh bā šādī wa salāmat.*

UA1 [MSP as the ML – Bandarī dialect as the EL]: *Qurbān-at kākā …*

CH7 [Bandarī dialect, then MSP as the ML – MSA as the EL]: *Inšallāh ki safar-i khūbī tūbūda ba-ša … wa khuš āmad ba-šumā ʿammī ʿazīz.*

---

10. UA1 typed the two sentences on an Arabic keyboard where he was unable to differentiate between the Persian voiceless velar stop /k/ and the voiced velar stop /g/ written with two strokes (گ). He used the Arabic /k/ in *saḥargāh* 'dawn' and *bar-gaštīm* 'we returned'.

UA1 [MSP as the ML – MSA as the EL]: *Khailī mamnūn ʿammī ʿazī khūb hasta jā-yi šumā khālī.*

UA1: 'We returned to the country today at dawn after one week of travelling == After one week of travelling, today at dawn we returned to Dubai.'

UA3: 'Welcome back.'

UA1: 'May God bless and keep you, [name].'

CH5: 'Welcome back.'

UA1: 'Thank you, dear uncle.'

CH6: 'Welcome, at your service, oh my brother [name]. As God said: Keep "Travel through the land" with you at all times along with happiness and good health.'

UA1: 'I'm your sacrifice, brother [name].'

CH7: 'God willing, you had a good trip. And welcome to you, my dear uncle.'

UA1: 'Thank you very much, my dear uncle, it was good, we missed you.'

The dialogue begins with an exchange in GCA, where UA1 types the affricated GCA form of the 2nd pers. fem. sing. object suffix *-ič* with a *šīn*: *isallim-iš* 'bless you', *ikhallī-š* 'keep you'. CH6's sentence is a representative example of bilingual-diglossic code-switching with MSP as the ML and MSA, Bandarī dialect as the ELs. The MSP substructure, *khuš āmadī … ṭibq-i farmān-i khudā … rā hamīša dāšta bāšī hamrāh bā šādī wa salāmat* 'Welcome … As God said … Keep it with you at all times along with happiness and good health', serves as a basis for intra-sentential code-switching with a lexical insertion from Bandarī dialect, *kāk* 'brother', syntactic insertions from MSA, *ʿalā ʿainī* 'at your service', *yā akhī* 'oh my brother', and an EL island from the Quran (29/20): *sīrū fī al-arḍi* 'travel through the land'. UA1's reply is a diglossic code-switching with MSP as the ML and a lexical insertion from Bandarī dialect, *kākā* 'brother'. CH7's post features a diglossic inter-sentential code-switching between Bandarī dialect and MSP, while the MSP sentence incorporates an adjectival construction from MSA with a possessive suffix, *ʿammī ʿazīz* 'my dear uncle'. This is an instance of how contact-induced language change results in a morphosyntactic borrowing. In MSA, the definite article *al-* is dropped from a nominal taking a possessive suffix, while it remains definite, and any qualitative adjective carries the definite article explicitly: *ʿammī al-ʿazīz*. The form attested in the conversation conforms to Persian morphological rules, where the definiteness of nouns and adjectives are unmarked. The fact that UA1, a native Arabic speaker, repeats the 'incorrect' Arabic form, i.e. the adjective without the definite article *ʿammī ʿazī*, suggests that this expression has become a lexicalized item.[11]

---

11. The letter /z/ missing from the end of the word in *ʿazīz* 'dear' is a typo.

(10) CHAF: The following conversation plays out among four friends from Kīsh. The dialogue is a series of congratulatory notes to KI3 on his upcoming marriage. KI3 is an Arabic-Persian bilingual speaker who replied to the posts directed at him in the language they were written in. The dialogue includes both inter- and intra-sentential code-switching with an attempt from KI5 to write his mixed Arabic-Persian sentence with the Latin script. KI2 and KI3 initially exchanged two comments in GCA, KI2 typing on a Persian keyboard (final form of the letter *kāf*: ک), and KI3 using an Arabic keyboard (final form of the letter *kāf*: ك). KI4's Persian message indicates how diglossic code-switching between the H and L varieties of Persian functions, while the response from KI3 is in CP with an insertion from Bandarī dialect. KI5's sentence is an intra-sentential code-switching between Bandarī dialect and GCA, to which KI3 answers in GCA.

KI2: يا هلا بالمستكبر

KI3: هلا بمستكبرنه اشلونك شخبارك شو امسوي

KI2: مبروك مبروك يالمستكبر

KI3: الله ايبارك بيك في انشاالله في كيش بعد شهر صفر

KI4: ماشاالله اقا تبريک ميگم انشاالله خوشبخت بشيد

KI3: قربونت گارداش انشاالله عروسي تو كيش بعد از ماه صفر

KI5: kerboon veld khaltii

KI3: هلا بيك ولد الخالة سلم علي خالتي وايد

KI2 [GCA]: *Yā halā bi l-mustakbar.*

KI3 [GCA]: *Halā bi-mustakbarna išlōn-ik š-akhbār-ik šū imsawwi?*

KI2 [GCA]: *Mabrūk mabrūk yā l-mustakbar.*

KI3 [GCA]: *Allāh ibārik bīk fī inšā Allāh fī Kīš baʿd šahr Ṣafar.*

KI4 [MSP as the ML – CP as the EL]: *Mā šā Allāh āqā tabrīk mī-gam inšā Allāh khušbakht ba-šīd.*

KI3 [CP as the ML – Bandarī dialect as the EL]: *Qurbūn-it gārdāš inšā Allāh ʿarūsī tū Kīš baʿd az māh-i Ṣafar.*

KI5 [Bandarī dialect – GCA]: kerboon veld khaltii. [= *Qirbūn wild khāltī.*]

KI3 [GCA]: *Halā bīk wild əl-khāla sallim ʿalā khāltī wāyid.*

KI2: 'Hello, Haughty One.'

KI3: 'Hello, our Haughty One, how are you? How are things? What are you up to? [= What are you doing?]'

KI2: 'Congratulations, congratulations, Haughty One.'

KI3: 'May God bless you. There is [a wedding], God willing, on Kīsh after the month of Safar.'

KI4: 'May God preserve you, Sir. I congratulate you, God willing you will become fortunate.'

KI3: 'Thank you [= I'm your sacrifice], brother, the wedding is on Kīsh after the month of Safar.'

KI5: '[I'm your] sacrifice, my cousin.'
KI3: 'Hello cousin, a lot of greetings to my aunt.'

The five Arabic sentences in the conversation reflect dialectal features of GCA in the use of interrogatives (*š-*, *šū* 'what', *išlōn* 'how'), verbal conjugation (*imsawwi* 'you are doing', *ibārik* 'blesses' with the initial *hamza* and *yā'*) and vocabulary (*hala* 'hello', *wāyid* 'very, many, a lot'). KI4 is a Persian-speaker, which necessitates the exchange between him and KI3 to occur in Persian. KI4 wrote in CP (*mī-gam* 'I say'), but the final word of the sentence, *ba-šīd* 'you become' is a hybrid form of MSP (*ba-šawīd*) and CP (*ba-šīn*): the MSP inflectional ending is added to the truncated CP verb stem. KI3's reply in CP is marked by the common vowel shift in *qurbān* > *qurbūn* 'sacrifice' and the use of CP *tu* 'in' (MSP *dar*), with a direct borrowing from the Bandarī dialect, *gārdāš* 'brother' (< from Turkish *kardeş* 'sibling'). Regarding code-switching, KI5's sentence is problematic. He used the Latin script to render (1) a Persian mannerism in Bandarī dialect (*kerboon*) and (2) an Arabic phrase in GCA (*wild khālti*). The sentence is inconclusive to assign the function of ML and EL to the expressions as either one can be interpreted as an EL island, while the other serves as the ML. Since KI3 counters the sentence with a post in GCA, I lean toward viewing GCA as the ML.

(11)   The speech of common members of the Āl ʿAlī is prone to incorporating a significant number of lexical borrowings from MSP/CP into GCA. These lexemes primarily come from the realm of fishing and shipbuilding, and but also from everyday activities.
CH8: *əl-markab šisma*? *musāfirkišī*. [GCA] 'The boat, what is its name? [CP] Passenger carrier.' *-kiš* is the present stem of the Persian verb *kišīdan* 'to carry' complemented with the abstract noun suffix *-ī*.
CH9: *tabbī sawwi gašta*? 'Do you want to take a tour?'

Persian lexeme *gašt* 'tour' with the Arabic feminine marker *-a*.

6. Conclusions

This paper demonstrated how code-switching involving Arabic and Persian is largely amenable to the analysis presented by the Matrix Language-Frame model. The coexistence of Arabic and Persian on Iran's Persian Gulf littoral and the shared cultural background of their speakers have allowed Iranian Arabs to undergo a varying degree of Persianization. Male members of the Āl ʿAlī's ruling family were more prone to intermarry with Persian-speaking women than ordinary tribesmen. This resulted in a dual Arabic-Persian identity among the last generation of sheikhs on the Iranian coast. When they moved to the Arabian Peninsula in their teenage

years, the Persian identity remained so strong that many of them later married Persian women from Iran. Their children are now Emirati or Qatari citizens receiving formal education in Arabic, while Persian is confined to their homes. However, average members of the tribe who still remain in Iran are less Persianized, and consider themselves exclusively as Arabs. The older generations who never attended school in Iran speak hardly any Persian, and need their children or grandchildren to serve as interpreters with local authorities. The younger generations are fluent in Persian, but rarely use the language when dealing with the local tribal community. They do not associate themselves with the Persian national and Shiite religious identity of the Islamic Republic. Instead, they feel that their culture and language are firmly rooted in Arabia. But when these young people move to the Arab Gulf States, they are required to retain their Iranian citizenship. Their children, who are born in the UAE or Qatar without much hope of acquiring local citizenship, are usually raised as Arabs in a linguistic and cultural sense.

The Āl ʿAlī elite and their fellow tribesmen in Chārak and Kīsh exhibit patterns of Arabic-Persian situational and transactional code-switching, controlled by components of the speech event such as topic and participants. It can be inter-sentential where the switching occurs at clause-boundaries, or intra-sentential with Embedded Language islands, typically lexical and syntactic insertions. The speakers generate a smoothly functioning tetra-glossic switching, common for people with roots on both coasts of the Gulf: MSA, MSP, GCA, and Bandarī/Ačumī dialects. Elements of contact-induced language change comprise phonological (insertion of MSA sounds into CP, and MSP/CP sounds into GCA) and morphosyntactic borrowing. No difference between the mechanisms of spoken and written code-switching was observed. The following variations occur in the speech of the Āl ʿAlī:

> MSP as the ML – CP as the EL
> MSP as the ML – Bandarī dialect as the EL
> MSP as the ML – MSA and Bandarī dialect as the ELs
> CP as the ML – Bandarī dialect as the EL
> Bandarī dialect as the EL – GCA as the ML, or vica versa
> MSP as the ML – MSA as the EL
> MSA as the ML – GCA as the EL
> GCA as the ML – MSP/CP as the EL

This research intends to contribute to the fields of interactional sociolinguistics, computer-mediated discourse analysis, and linguistic anthropology, in addition to opening up new horizons in discussions about ethnic, cultural and religious identity in a region where Arabs and Persians share a vibrant history. The language data I recorded and continue to record in the Gulf is the first step in shedding light on the extent of the remaining Arab culture in Iran.

7.   Bibliography

Al-Anṣārī, Jalāl Khālid (2014). Qabāʾil ʿArab al-Hawla fī ʿUyūn al-Raḥḥāla wa al-Mustašriqīn. *Bairūt: al-Dār al-ʿArabīya lil-Mausūʿāt.*

Al-Dailami, Ahmed (2014). 'Purity and Confusion': The Hawala between Persians and Arabs in the Contemporary Gulf. In Lawrence G. Potter (ed.), *The Persian Gulf in Modern Times: People, Ports, and History* (pp. 299–326). New York, NY: Palgrave Macmillan.

Bakhtiārī, Majīd (2001). *Rāhnamā-yi mufaṣṣal-i Īrān: Ustān-i Hurmuzgān*. Tihrān: Gītāšināsī.

Floor, Willem (2014). *The Hula Arabs of the Shibkuh Coast of Iran*. Washington DC: Mage Publishers.

Gazsi, Dénes (2008). Arabic Proverbs from Iran: Some Characteristics of Arabic Spoken in Khūzistān Province. In Stephan Procházka, Veronika Ritt-Benmimoun (eds.), *Between the Atlantic and Indian Oceans, Studies on Contemporary Arabic Dialects* (pp. 195–205). Wien & Münster: Lit Verlag.

Lorimer, John Gordon (1970). *Gazetteer of the Persian Gulf, 'Omān, and Central Arabia*, Volume II A & II B. Westmead: Gregg International Publishers.

Myers-Scotton, Carol (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198299530.001.0001

Nadjmabadi, Shahnaz R. (2009). The Arab Presence on the Iranian Coast of the Persian Gulf. In Lawrence G. Potter (ed.), *The Persian Gulf in History* (pp. 129–145). New York, NY: Palgrave Macmillan.

Nūrbakhš, Ḥusain (2003). *Banādir-i Īrān dar Khalīj-i Fārs*. Tihrān: Daftar-i Pažūhišgāh-i Farhangī.

Pelevin, Mikhail (2010). Materials on the Bandari Dialect. *Iran and the Caucasus* 14, 57–78. https://doi.org/10.1163/157338410X12743419189388

Poplack, Shana, Sankoff, David, Miller, Christopher (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26(1), 47–104. https://doi.org/10.1515/ling.1988.26.1.47

Potts, Daniel T. (2004). Kish Island. In: Ehsan Yarshater (ed.), *Encyclopædia Iranica*, online edition, available at http://www.iranicaonline.org/articles/kish-island

Saršumārī-yi ʿumūmī-yi nufūs wa maskan (2011). Available on amar.org.ir

Sayahi, Lotfi (2014). *Diglossia and language contact: language variation and change in North Africa*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139035576

Schmitt, Rüdiger (1989). *Compendium Linguarum Iranicarum*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Walters, Keith (1996). Diglossia, linguistic variation, and language change in Arabic. In Mushira Eid (ed.), *Perspectives on Arabic Linguistics VIII: Papers from the eighth Annual Symposium on Arabic Linguistics* (pp. 157–197). Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.134.12wal

Arabic borrowing of the Hebrew word *menahēl* 'manager'
Articulations and ideologies

Nancy Hawker
University of Oxford

Ideologies, or ways of understanding one's relation to the world, impede or encourage, and affect the form of, language contact practices such as borrowing and codeswitching. This is illustrated by the pragmatic functions – informative or humorous – of the Israeli Hebrew word *menahēl* 'boss' in Palestinian Arabic. By using 'boss' in an ironic sense, to refer to a self-important 'big-head', Palestinians are expressing their stance by means of a Hebrew loanword, to take a dig at the powers that be. The article provides examples of real usage and grounds the explanation for the different meanings in pragmatics, cultural theory, and Althusser's conception of ideologies in ways that are useful to linguistic ethnography.

Keywords: Palestine, Israel, migrant workers, borrowing, codeswitching, humour, ideology, political economy, pragmatics, Arabic, Hebrew, linguistic ethnography

1. Introduction

This chapter examines Arabic usages of the Hebrew word *menahēl* which means 'manager', 'foreman' or 'boss' when spoken by modern Israeli Hebrew speakers (Levy 1995: p. 175). This case of borrowing illustrates two patterns of linguistic practices within the field of Palestinian Arabic/Modern Hebrew contact. These patterns are linked to ways of thinking about the relationship, including relations of conflict and economic inequality, between the groups of speakers identified with these languages. The explanation for patterns of uses of Hebrew borrowings in Arabic will contribute to two intellectual projects of linguistic anthropology and ethnography by demonstrating, firstly, the role ideologies play in engendering language change (Kroskrity 2000), and, secondly, the usefulness of a social analysis of

language in gauging power-inflected processes at work in societies (Rampton 2009). The framing of the contribution furthermore offers a characterisation of *ideologies* valuable for the examination of language in society, and a conceptualisation of the link between social practices, including linguistic practices, and ideologies, in the form of *articulations* (Hall 1980, 1996).

Usually, *menahēl* is borrowed by Palestinians to simply mean the manager of the workplace as an equivalent, for instance, to the Palestinian Arabic *maʿallim* 'foreman'. However, evidence from the occupied West Bank (Hawker 2013: pp. 90–2) and from the Egyptian Sinai (Holes & Abu Athera 2009: p. 15) shows that it can also be used by Arabic speakers to sarcastically denote a 'big-head' with connotations of misplaced authority and self-aggrandisement: a meaning that does not exist in the original Hebrew. The traces of the power relations in the Palestinian-Israeli context are evident in *menahēl*: there is the economics – it means 'boss' – and there is the politics – it comes from Israeli Hebrew, indexing the national identity of the dominant group. The explicitness of the context in the particular lexical item we are focusing on makes it imperative to integrate the context into the analysis, while acknowledging that the contexts are, at their particular and changing historical junctures, always mediated by ways of thinking about them: what is termed here *ideologies*.

Menahēl is just one example, albeit an interesting and poignant one, of two broad patterns of pragmatic functions of Hebrew loanwords. One, the pattern of interchangeable use of Arabic and Hebrew equivalents, whereby the Hebrew is minimised in the presence of out-group interlocutors, is associated with the speech of Palestinian day-migrant workers employed inside Israel or in the occupied West Bank's Israeli settlements, illegal under international law. The second pattern is 'ironic power humour', the witty narrative manipulation of Hebrew's association with Israeli power by using the Hebrew loanwords or codeswitched phrases in paradoxical situations. The identification of these patterns is based on the analysis of material (going beyond *menahēl*) collected during fieldwork in the West Bank in 2007–8 and in Israel in 2015 and can be applied also to material collected by other researchers.

The explanatory model links, or *articulates*, particular instances of language use with a pattern of linguistic practices, and an *ideology*, developed for understanding the experienced context which is captured in the language. There are more than two patterns and ideologies operative in directing Arabic speech towards the use or avoidance of Hebrew borrowings and codeswitching; only those relevant to the examples provided will be sketched out in this chapter. The chapter will start with six examples of Arabic usage of words for 'boss'. The sorting of the instances of *menahēl* into patterns will be supported by additional, selected, empirical evidence that will show the spread of the identified pragmatic functions. On the basis of the

empirical linguistic material, and of knowledge of the Palestinian-Israeli contexts relevant to the language practices, the explanatory model involving *articulations* and *ideologies* will be laid out. In conclusion, the theoretical implications of *articulation* and *ideology*, including proposed definitions of these terms, will be offered as contributions not only to the fields of language contact research and linguistic anthropology and ethnography, but also to discourse analysis, political thought and Middle East studies.

2.   Who's the 'boss'? *Maʿallim* (Arabic) and *menahēl* (Hebrew loanword), and the speech of Palestinian day-migrant workers

During peak years of labour migration to Israel after the 1967 military occupation of the West Bank and Gaza Strip, some 45 per cent of Gazan workers and nearly 30 percent of the West Bank workforce were employed in Israel (Farsakh 2005). In some refugee camps, 80 per cent of male refugees of employable age worked in Israel (Farsakh 2005). Palestine refugees were particularly disposed to becoming a floating source of manual labour since the vast majority of them had been peasants in pre-1948 Palestine and thus lost access to the land that had been the source of their livelihood (Pappé 2006).

The numbers of Palestinian migrant workers are now reduced due to a securitised regime incrementally formed from the beginning of the 1990s, involving periodic closures causing intermittent high unemployment, and an opaque system of travel and work permits, which, though justified officially on security grounds, grants, for instance, easier access to Israeli settlements in the West Bank than to Israel (Farsakh 2005). The restrictions are by-passed by workers who have few other employment options, and therefore accurate current figures on employment of Palestinians in Israel are hard to come by, though they are estimated at around 110,000 migrant workers out of a West Bank workforce of 810,300 (Palestinian Central Bureau of Statistics 2015).

The remittance of wages earned in Israel had, in the past, been a significant source of Palestinian Gross Domestic Product increase, if this can be measured for a territory without defined borders (Fishelson 1992), though this labour had not contributed structurally to the development of the Palestinian economy. Rather, this labour was seen to enhance the standard of living of individual workers, who invested in family housing, but who also experienced reduced expectations of education and skilled training, especially for men, who were then unequipped to adapt when economic shocks hit the Palestinian workforce as a result of Israel's securitisation (Roy 1999).

Though employment in Israel continues to be a source of contact of Arabic speakers with Hebrew, this contact is inflected not only by economic considerations, giving rise for instance to a borrowed lexicon for tools, but also by the securitisation that dominates their access to work, reflected in terms relating to military control. Moreover, though employment in Israel, when it is possible, is a source of relatively high wages, it entails putting up with poor health and safety conditions and job precarity (Lewin-Epstein & Semyonov 1987; Alenat 2010). These aspects of the context, and more, and the ways Palestinians think about them, are reflected in the instances of speech recorded in the following four examples. The speech is transcribed using academic Semitic orthography and Hebrew loanwords are in italics both in the original Arabic sentences and the English translations. Loanwords from other languages are in italics and underlined. The terms for 'boss', *maʿallim* or *menahēl*, are in bold.

> Example (1)
> Palestinian speech using *maʿallim* 'foreman' (from a 2008 interview with a Palestinian man in Dheisheh Refugee Camp)
> marra kān nfič fi ḥēṭ iṭ-ṭubār. ma manṣūba is-*skāla* ʿala li-ḥbāl … ṭābik iṭ-ṯāliṯ, ana u uḫra ʿāmil. w-eḥna nfič fi l-ḥēṭa fi *fībar*, haftat is-*skāla* aža bidha tikaʿ min šakti u min šakta ṯābta. fa hū masaʿ il-ḥabil, ana masačit fi l-*kutsīm* ḥadīd sitta mīli ṯāliʿ min il-ḥēṭ. masačit fī u hū nafs il-'iši masaʿ fīhen u ṭleʿna ʿa l-ḥēṭ fa šāfna il-**maʿallim**.
> 'Once, we were stripping the formwork from a wall. The *platform* wasn't hanging on the ropes … We were at the height of the third floor, me and another worker. While we were stripping from the wall with the *grinder*, the *platform* slipped and was just about to fall on my side, but was steady on his side. So he grabbed the rope, I grabbed the six millimetre-strong metal *spikes* that were sticking out of the wall. I held on to them, and he did the same, he held on to them, and we climbed to the top of the wall, and the **foreman** saw us.'

The speech exhibits features typical of the rural Palestinian register, known as the *fellāḥi* 'peasant' variety, including the /č/ in *nfič* 'cutting, stripping' and *masačit* 'I held, I grabbed' , the urban equivalent of which would be pronounced with /k/. This is significant in several sociolinguistic ways: we know the speaker to have been born and raised in Dheisheh Refugee Camp in the West Bank near Bethlehem. Despite his proximity to the Palestinian urban centre, where the Jerusalem urban dialect is the prestigious norm, he retains *fellāḥi* speech features, betraying the rural roots of his family's background in pre-1948 Palestine. His speech also indexes his position in the political economy: it can be inferred that the speaker is educated at a basic level, due to the absence of a formal register in the interview context where more

educated speakers would borrow from prestigious urban or Standard Arabic, and that this absence is also a consequence of social, gendered, expectations of young men. Young Palestinian women from equivalent socio-economic groups would be socially expected, in general, to receive higher levels of formal education (Jacobsen 2004), giving them repertoires of a relatively more formal register commonly indexed in an interview.

The technical terminology is referred to casually in Example (1), without explanation provided to the interviewer who is not a construction worker, in the breathless account of a dangerous incident which is the focus of the narrative. The focus is an indictment of the poor safety provisions as well as the anxious relationship with the *maʿallim* 'foreman' who is a Jewish Israeli, because this incident occurred on a construction site on Jabal Abu Ghneim where the Israeli settlement of Har Homa has been built since the 1990s. Several specialised terms for technical items pass unexplained, because their precise meanings are peripheral to the thrust of the narrative, such as *ṭubār* 'formwork' or *kutsīm* 'spikes', from the Hebrew *kotsim*, which are the metal bars used in the casting of reinforced concrete. Other technical terms in the speech, *skāla* (plur. *sakāyil*), 'construction platform', and *fībar*, 'disc grinder', are also not of Arabic origin and are found in Arabic speech in the region of the East Meditarranean. *Skāla* originates in the Italian for 'ladder', *scala*. The spread of Italian loanwords in this type of speech is another indicator of the political economy's role in spreading language change, since it has been proposed that the source of this vocabulary were migrant workers from southern Italy in the early 20th century (D'Anna 2018, this volume). The flows of movement of the manual workforce has been a feature of the industrialisation of the Mediterranean space (Beinin 2001) which affected borrowing for vocabulary related to that type of work whereas vocabulary related to agriculture remained relatively stable.

> Example (2)
> Palestinian speech using the plural of *menahēl*: *manahīl* 'bosses' (from a 2008 interview with a Palestinian man near Tulkarem)
> NH: kīf ir-rātib?
> 'How's the pay?'
> Worker: wallāhi, hassa māši. hassa.
> 'Well, now it's ok. Now it is.'
> NH: kān aswa?
> 'It was worse?'
> Worker: kān taḥt il-ḥadd il-ʾadna. fa aḍrabna. mā smeʿtūš?
> 'It was below the minimum wage. So we went on strike. Haven't you heard?'
> NH: fa fuztu.
> 'And you won.'

> Worker: ā. fa tadaḫḫalat il-*lēška* u ṭalabat rātib qānūni. fi l-mufāwaḍāt maʕ il-**manahīl**.
> 'Yup. And the *bureau* intervened and demanded legal wages. In the negotiations with the **bosses**.'

The incident the worker is referring to, an industrial strike at an Israeli settlement-based factory known as Geshuri near Tulkarem in the northern West Bank in autumn 2007, would have been known to local residents but the presence of a foreign interviewer served as a pretext to briefly recount the events for which Hebrew loanwords were usefully descriptive. Palestinian employment in Israeli settlements is not supported by Palestinian national institutions (and indeed has been made illegal since 2010) because the settlement project contravenes international law, generally furthers the seizure of Palestinian natural resources, and leads to a myriad of human rights violations. At the same time, Palestinian workers are considered to be outside the jurisdiction of Israeli labour laws. This situation was changing during the course of the late 2000s under pressure from workers' actions supported by an Israeli non-governmental organisation, Kav LaOved, which raised awareness of workers' rights in Arabic (Alenat 2010). The self-organised Palestinian workers in Geshuri learnt specialised vocabulary from Standard Arabic such as *il-ḥadd il-ʾadna* 'minimum wage' in Arabic as evidenced in Example (2), yet the institution that finally intervened in their favour was the 'bureau', pronounced *lišká* by Israelis (the acute accent indicates stress) and *lēška* by Palestinians. *Il-lēška* 'the bureau' (the Hebrew loanword is preceded by the Arabic definite article) refers to a body at the Israeli Ministry of Industry, Trade and Labour which is responsible for the enforcement of workers' rights according to Israeli laws. The Hebrew loanword here is representative of power relations that preclude Palestinian institutional support for workers employed by Israelis due to nationalist principles which have nevertheless failed to deliver a viable economic alternative for the workers.

Similarly to the most common loanword from Modern Hebrew into Palestinian Arabic, *maḥsōm* or *maḥsūm* 'checkpoint', *menahēl* has Arabic plural patterns applied to it: whereas the Hebrew plural of *menahel* (pronounced in Israeli Hebrew with a short vowel) is *menahelim*, and of *maḥsom* is *maḥsomim*, in Arabic the plurals follow one of the patterns for broken plurals to form *manahīl* and *maḥasīm* respectively.

> Example (3)
> Palestinian speech using the feminine plural of *menahēl*: *menahelāt* 'women managers' (from a 2008 interview with a Palestinian woman in Tulkarem Refugee Camp)
> Worker-Cleaner: štaġalt sitte snīn *nekayōn*.
> 'I worked as a *cleaner* for six years.'

NH: fi dūr in-nās, willa …?

'In people's houses or …?'

Worker-Cleaner: fi mustašfa aʿṣāb, yaʿni, ʿind il-maǧanīn. šuǧul mrattab. aṭlaʿ ʿala sitte u ʿala waḥde akūn fi dāri. u fi šahər āḫod alfēn šēkel.

'At a psychiatric hospital, I mean, where the crazy people are. It was a well-organised job. I'd leave the house at six and by one I'd be back. And in one month I'd get two thousand shekels.'

NH: mā kuntīš itḫāfi min il-marḍa?

'Weren't you afraid of the patients?'

Worker-Cleaner: laʾ, humme baku yinaḍfu ǧurfethum laḥālhum … bištiǧlu, yaʿni, kulši, byuḍrubūš ḥada. ana bakēt anaḍḍef bas ʿind il-***menahelāt***, fi l-makātib. il-***menahelāt***, yaʿni, zayyik.

'No, they cleaned their rooms themselves. They work, I mean, [they do] every-thing; they don't hit anyone. I was cleaning just the ***women managers'*** offices. The ***women managers*** are, I guess, like you.'

Describing one's job by using a Hebrew loanword, as *nekayōn* 'cleaning' or *binyan* 'construction', is typical for Palestinian day-migrant workers, though they also resort to the Arabic equivalents *tanḍīf* and *buna* respectively (Hawker 2013: pp. 34–66). However, some usages indicate that the Hebrew loanwords might be semantically more restricted as it only refers to the jobs of this kind in Israel, not in general. *Nekayōn* can never be used to describe cleaning one's own home as a domestic chore, whereas *tanḍīf* can.

The speaker from Tulkarem Refugee Camp who spoke favourably of Israeli women managers and of psychiatric patients in Example (3) took up waged em-ployment in Israel because the men in her family were either dead or in prison. The researcher's question regarding fear of the patients was motivated by a search for topics that would trigger affective speech or a longer stretch of narration from the interviewee in order to record the language practices resulting from contact with Hebrew, and does not reflect any genuine position toward psychiatry. The cleaner stopped working at the hospital in Israel because of the movement restric-tions imposed by the Israeli army in the Tulkarem area after 2002. Her positive opinion of the 'well-organised' cleaning job is based on the reported facts that she worked hours that suited her family commitments, and was paid regularly. What this worker could not abide, and what the Palestinian migrant workers complain of generally, is the securitisation of the access to work.

Example (4)

Imagining a pleasant commute to work (from an interview in 2006 near Bethlehem)

bas lāzem yikūn ashal, il-murūr. kāwnu yaʿṭīk *te'uda* min iš-šurta innak inte … fiš ʿindak wala *nikuda*. inte ʿindak ṭaṣrīḥ li l-ʿamal fi israʾīl, ōke, ḫalaṣ,

šū il-muškile? fiš muškile. bišūf iṭ-ṭaṣrīḫ, yišūf il-bani ādam, il-hawīye tabʿato, *bevakaša*, tfaḍḍal.

'But it should be easier, the transit. If they give you an *ID card*, from the police, [certifying] that you don't have any *penalty points*, you have a permit to work in Israel, then *OK*, enough, what's the problem? There is no problem. He [an imagined helpful soldier at a checkpoint] sees the permit, sees the person, his ID card, *here you go*, welcome.'

Hebrew loanwords for military bureaucratic terms such as 'ID cards' and 'penalty points' (which proscribe the awarding of work permits), and for the politeness formula *bevakaša* accompanying a fictional pleasant commute through an army checkpoint, are only to be expected in a complaint about the problems faced by workers experiencing collective restrictions on freedom of movement. Accepted precarity, appreciation of a low but regular wage, and demand for efficient access: these standards viewed as normal by the Palestinian migrant workers perhaps make them the ideal subjects of neoliberal economic models coupled with pervasive securitisation.

Examples (1)–(4) do not comprehensively treat the language practices of Palestinian migrant workers in Israel (for more, see Hawker 2013). It might seem that the context is given disproportionate weight in the descriptions of the speech excerpts. Furthermore, the excerpts themselves are relatively long, allowing the respondents to provide an account of the relations they experience and how they regard them in their own words. The contention of this chapter is that these elements – the context, and the stance of the speakers – are central to the analysis of even one loanword, *menahēl*. What is notable at this stage of the analysis is that the borrowed vocabulary has Arabic equivalents or approximations which can be used interchangeably or as glosses: *menahēl* can be substituted by *maʿallim* without any loss of meaning, *te'uda* 'ID card' is also referred to by the Arabic *hawīye* in Example (4), *nekayōn* 'cleaning' is *tanḍīf*, and even the *kutsīm* 'spikes' in Example (1) is described in such detail that an uninitiated interlocutor can understand without knowing the word that it is a thin metal element like a rod that protrudes from the wall and can be grasped by a human hand. The uninitiated interlocutor, in this case the researcher, is a factor in the analysis, representing the out-group with no experience of work in Israel, who requires translations and explanations in the immediacy of the conversation that are not necessary in communication among co-workers, and who may also, in the wider context, be aware of the nationalist principles and ideas of economic developmentalism that take a negative view of employment outside the Palestinian sector. This factor contributes to the de-emphasising of Hebrew loanwords in the speech of Palestinian day-migrant workers. This has to be contrasted with the perspective emanating from the workers own words – the necessity of waged work, the different modulations of power relations with the

bosses, and the difficult yet unavoidable securitised access – which pulls the speech practices towards some use of Hebrew.

The only Hebrew item in the preceding examples which would lose some of its meaning if it were rendered only in Arabic is *bevakaša* 'here you go, welcome', itself glossed by the Arabic *tfaḍḍal*, in Example (4). The sentence 'he sees the permit, sees the person, his ID card, *here you go*, welcome' conjures the improbable image of a helpful polite soldier at a checkpoint, in contrast with the workers' lived experience of receiving hostile orders. The incongruity of a soldier welcoming the Palestinian on his way to work with *bevakaša*, in this account, is an element of narrative humour: *tfaḍḍal* on its own would not have conveyed the irony, which brings us to the second pattern of use of *menahēl*.

3.   No kudos for the 'big boss': *Menahēl* in ironic power humour

One woman used *menahēl* in conversation to describe the baseness of the behaviour of her brother-in-law who had had the impertinence of snubbing her sister (his wife) by marrying several wives in addition to her.

> Example (5)
> The polygamous 'big-head' (from a 2007 conversation with a woman in Shuafat Refugee Camp)
> tǧawwaz tintēn ʿalēha. u hūwe ustāz fi l-ǧamiʿa! hūwe il-**menahēl** li-kbīr.
> 'He married two women on top of her. And he's a university professor! He thinks he's the **big cheese**.' (Literally: 'He is the big **boss**.')

The speaker is mocking the man by contrasting his status of university professor with the uncouth marital polygamy, and she has at her disposal one word that captures this irony: *menahēl*. The English gloss is not as concise, requiring several approximations in Example (5), unless one resorted to vulgarisms which would adequately convey the expressed stance towards polygamy. Palestinian native speakers of Arabic, when tested, have similarly struggled when asked to render this meaning with Arabic words. The reason why this is possible to do specifically with the Hebrew loanword relates transparently to the wider social, economic, political and indeed military power relations in the region. Taking the word for 'director' in the language of the dominant state, and then subverting its meaning to imply illegitimate, self-aggrandising authority, is a way of having a dig at the powers that be, even if this is the power of a man to engage in sexual behaviour forbidden to women.

A more explicitly political use of *menahēl* was recorded in a village near Nablus in the northern West Bank which had been demolished by the Israeli army in January 2010. The speaker, an elderly woman, was a shepherd and cheese-maker.

The setting was a tent of plastic sheeting that had replaced her demolished dwelling, and the audience was primarily a group of younger women whom she was entertaining, though the presence of a foreign human rights researcher accompanied by the mayor who deferred to the speaker's seniority had prompted this particular gathering.

> Example (6)
> The demolition 'gangmasters' (from an interview in 2010 with a woman in a village near Nablus)
> iǧǧat iǧ-ǧarrafāt u ana bakēt aʕmil fi ǧ-ǧibne. il-ʕummāl u ǧ-ǧunūd u l-**manahīl** tabāʾhum wikfu ḥawaley. ḥačēt ilhum šū biddču, nitfet ǧibne?
> 'When the bulldozers came I was making cheese. The workers, the soldiers and their **so-called bosses** stood all around me. I told them, "What do you want, a bite of cheese?"'

The speech was delivered in the *fellāḥi* 'peasant' variety similar to Example (1) but here it is sociolinguistically unremarkable since the speaker was indeed a villager and an older woman with very little formal education, unlike the young male construction worker living in the Bethlehem urban agglomeration. Another difference is that in Example (6) the speaker had the charisma to rhetorically dominate the encounter and thus change the applicable linguistic norms: while an interview geared at obtaining a factual report might be expected to elicit short informative sentences in a relatively more formal register at least initially, the narrative performance in Example (6) entertained an audience of fellow villagers. The account of the demolition was given in derisive tones and the speaker's snubbing of *il-manahīl* 'the so-called bosses' was met with laughter. Using Arabic equivalents such as *ḍubbāṭ* 'officers', or *mudarāʾ* 'directors' would have been factual, not funny, and the insertion of 'so-called' in the English gloss only approximates the ironic and irreverent connotations.

There is evidence that *menahēl* in its ironic sense has spread beyond direct Palestinian-Israeli contact in the use by a poet from the Sinai, Ḥusayn bin ʿĪd bin Ḥamad bin Mişliḥ bin ʿĀmir al-Ṭayāḥā.

> Example (7)
> 'foreign clever-dicks' make good cars (from Abu Athera and Holes' collection of Bedouin political poems)
> wi l-*mārka* wi l-isim gālu tyūtah
> ma yiḥill bak ṣanʕat **manāhīl** *ḫēḫām*
> 'Toyota is her model and Toyota is her *make*,
> The workmanship won't disappoint: these **foreigners** don't fake!' (Holes & Abu Athera 2009: 15)

Ḥaḥam in Israeli Hebrew means 'clever'. Holes and Abu Athera explain the use of *manāhīl ḥaḥām* in this poem, which praises the virtues of a powerful car similarly to the traditional subject of racing camels, as pointing to the foreign-ness and skilfulness, and ultimately the craftiness, of the manufacturers. Anyone who is not Bedouin is a priori suspect: the Japanese car-makers are, to quote the translators, 'foreign clever-dicks', and a sarcastic Hebrew loanword about the 'bosses' is a good way to express that.

> Example (8)
> Palestinian humour invoking the Oslo peace process (from a school yard in Shuafat Refugee Camp in 2007)
> Teacher: taʿāl sallim ʿaleyy bi-l-ʾīd. ʿimilna ḥudna. btaʿraf šū hāda yaʿni? fi šū bitzakrak hādi ṣūra? zeyy rabīn u abu ʿammār fi-l-bēt il-abyaḍ. *šalom amiti.*
> 'Come shake hands with me. We've implemented a ceasefire. Do you know what that means? What does this picture remind you of? Like [Yitzhak] Rabin and Yasir Arafat at the White House. *True peace.*'

The teacher seemed accustomed to regulating school yard behaviour, and was resigned to the inevitable fights and short-lived 'ceasefires'. Yet he used his sense of humour to cope with two situations: the stress of acting as the arm of justice among young children, and the hollowness of the promise of 'true peace' pictured in the Rabin-Arafat handshake. To capture this irony he could use the Hebrew phrase *šalom amiti* 'true peace' – a political slogan of the 1990s – in this paradoxical situation.

Humour here serves the function of a coping mechanism in circumstances of adversity. Some types of humour permit the suspension of conventions of politeness (Barbe 1995: p. 89). Israelis, or generally some sort of despised authority, indexed by the use of Hebrew, are either directly or indirectly the butt of the jokes. Arabic cannot be a substitute in these situations because it does not index the relations of power that the Hebrew can. Moreover, using Hebrew for sarcastic humour not only suspends norms of politeness – it is not as vulgar as an expletive and not as rude as a command – it also suspends norms of language use. The Palestinian-Israeli conflict contributes to nationalist norms that regard Hebrew 'interference' as unpatriotic, in accordance with the language purism project that accompanied the promotion of Standard Arabic as a formal register. However, when Hebrew codeswitching and borrowing is employed to subvert the Israeli Hebrew meaning of 'boss', or poke fun at some stereotype of Israeliness, nationalist language purists face a quandary, because the function of using Hebrew in this way is actually aligned with nationalist ideology (for more examples of this type of humour see Henkin 2009; and Hawker 2013).

Humour is sometimes depicted as an outlet for resisting oppressive political power (Hodge & Mansfield 1985: p. 197). However, the irreverence of humour can

also be tolerated by an oppressive power as a sanctioned limited arena for criticism that does not spill over into political change (Eagleton 1981). It is not the Israeli state that would find the humorous use of Hebrew subversive, if it recognised it; rather it would be those Palestinians who adhere to nationalist norms that frown on extensive borrowing from Hebrew, and this is somewhat ironic.

4.   An explanatory model for the uses of *menahēl*: Articulations and ideologies

The analysis of particular instances of Arabic uses of *menahēl*, from the Israeli Hebrew for 'boss, manager, foreman', allows us to identify two broad patterns distinguished by pragmatic function. One (seen in Examples 1–4) is associated with the speech of Palestinian day-migrant workers and involves minimising Hebrew loanwords in interactions with Arabic-speakers who have no experience of work in Israel thanks to the possibility of substituting or glossing in Arabic. The other pattern (in Examples 5–8) can be used by anyone with the rhetorical skills to achieve the balance of humour and irreverence that are elements of irony and sarcasm, conveyed by deploying Hebrew for its connotations of Israeli power in paradoxical situations. For this function, Arabic cannot provide a succinct substitute.

In turn, the patterns thus summarised can be linked – or, *articulated* with – *ideologies* that mediate the experienced context. In the Palestinian-Israeli context, several ideologies are at play, and I give them descriptive labels for convenience. The Palestinian migrant workers' stoicism, or pragmatism (sometimes termed 're-silience'), is the rationalisation of what has to be done to make a living under the Israeli occupation that shapes the Palestinian economy and other aspects of daily life. This way of thinking is to some degree shared by all Palestinians who carry an Israeli ID card, since it is necessary to get by, and the stoicism negotiates in specific ways with other ideologies: Israeli securitism that puts Israel's military objectives ahead of any other consideration, with the neo-liberal ideology that cheap unprotected labour is an economic asset, and with Palestinian nationalism that demands collective resistance and sacrifice.

The explanatory model posits that ideologies serve as tractors that are articulated with, by means of a metaphorical flexible hinge, one or more trailers which are social practices viewed in analytical categories, for example, linguistic practices, and can be thought of together (Hall 1980). In this case, securitism and economic neo-liberalism are hegemonic global ideologies which the Palestinian leadership itself has espoused at least during the 1990s in the form of the Oslo process, and to which there seem to be only marginal voiced alternatives (Turner & Shweiki 2014). Put crudely, *maḥsūm* 'checkpoint' and *menahēl* 'boss' are loanwords from Israeli

Hebrew into Palestinian Arabic that are drawn into use by securitism and economic neo-liberalism. The flexibility of articulation, however, allows for the assignment of the same practices to another tractor, Palestinian stoicism, which acknowledges the negotiation with yet another ideology, Palestinian nationalism. Symbolic condemnation of work in Israeli settlements, denying institutional support for the workers' rights of Palestinian labour migrants, ineffectual official protest against the excesses of Israeli military operations, and disapproval of the use of loanwords from Hebrew, are articulated with Palestinian nationalism. Stoicism combines all these 'top-down' ideologies with comprehending how their contradictions play out in daily lived experiences. The outcome of the negotiation is a linguistic practice that incorporates loanwords but minimises their salience (thus indexing the contradictions) by substituting or glossing with Arabic in interactions outside of an in-group of migrant workers.

Palestinian nationalism and stoicism are negotiated with Israeli securitism differently when articulated with the linguistic practice of 'ironic power humour'. The success of the humour depends on the audience's shared experience of Israeli securitism, which the linguistic practice references, and also critiques, in paradoxical situations. The critique articulates with Palestinian nationalism, and the humour is a coping mechanism articulated with stoicism.

The analytical categories of the tractor-articulation-trailer metaphor allow to take instances of speech containing borrowing (the linguistic practice under examination) which are ordered according to pragmatic functions, link these to social practices (coping with living under Israeli control, being disappointed by the failure of the Oslo process, working in conditions of precarity and securitisation, organising the social roles of gender), and link these to ideologies through a series of articulations. Nothing is clear-cut, the articulations are flexible rather than deterministic, yet the model provides a nuanced explanation that incorporates relations of power made intelligible through ideologies which pull the practices in certain directions. These nuances are needed in the field of Hebrew-Arabic language contact so as to move beyond explanations relying on national identities and conflict, which are categories produced by nationalisms and as such explain only one facet of the language contact phenomena.

5. Conclusion: *Articulations* and *ideologies*: Definitions and links to broader questions

Work done in the field of linguistic anthropology and ethnography recognises the role of language ideologies in shaping the contours of the range of linguistic practices (Bucholtz & Hall 2005; Kroskrity 2000). Discourse analysis, pragmatics,

sociolinguistics and conversation analysis have developed techniques to isolate specific linguistic phenomena in actually-existing linguistic practices that are particularly sensitive to the effects of language ideologies (Hasan 2004). The social recognition of the ideological sensitivity of a phenomenon, such as Hebrew borrowing into Arabic, is *enregisterment* (Agha 2005), and the expression of an attitude toward such a phenomenon is *stance* (Jaffé 2009; Du Bois 2007). The explanatory model applied here to the case of *menahēl* 'boss' and other Israeli Hebrew borrowing or codeswitching in Arabic relies on this scholarship and has integrated these insights into the analysis in the preceding sections of this chapter. At this point in the chapter we broaden the scope by exploring what exactly is meant by *ideologies* and *articulations*.

The explanatory model posits that language ideologies are an integral part of ideologies generally, as has been demonstrated for instance in relation to monolingualism and nationalism (Silverstein 2000). What requires unpacking is the functioning of ideologies in the historically specific settings of the Hebrew-Arabic language contact at hand. A way of unpacking this that can be generalised to other settings would demonstrate the unique utility of analysing language practices for understanding other social practices in articulation with them and with ideologies (Rampton 2009).

What might not have been clear so far in the chapter is the concept of the tractor (ideology) and its articulations (configurations) with the trailers (practices). The tractor does not have a driver, for instance a powerful political group, such as Israeli military advisors 'designing' securitism. Nor is the tractor driverless, an (ideal) idea that can cause practices to occur, for instance to inspire people to avoid using Israeli Hebrew terms in conversations with uninitiated interlocutors. Nor is the tractor an 'outlook' that colours (distorts) in the minds of those who are looking (which is everybody) at another, separate ('alienated'), entity called 'the world'. We are not talking of ideology as a 'world outlook' then, but rather offering the tractor as a spatial and vectoral representation of the 'materiality of ideology'. With this oxymoron, Louis Althusser (1971: pp. 155–156) sought to capture what ethnographers of language know by necessity: that *ideologies* are only ever found in *practices*. Ideologies are performed (Butler 1997), do not exist independently of that performance, and shape the content of the performance: so that for instance both the soldiers' orders to present an ID card for inspection at a checkpoint, and the compliance with that order, are performances of securitism, which therefore exists materially in the shape of these practices. What we are saying here is not that orders and compliance are the same, neutralising relations of power, but that there is a common logic to both, and that logic pulls (like a tractor) the actors to perform in this way, in accordance with their ideas of the world, as observed ethnographically. This conception allows Althusser to arrive at the following definition of ideology:

"What is represented in ideology is therefore not the system of the real relations which govern the existence of individuals, but the imaginary relation of those individuals to the real relations in which they live" (1971: p. 155). And Althusser elaborates: "'Individuals' […] live in ideology, i.e. in a determinate (religious, ethical, etc.) representation of the world [which] depends on their imaginary relation to their conditions of existence" (1971: p. 156). And furthermore: "The ideology of ideology […] recognises, despite its imaginary distortion, that the 'ideas' of a human subject exist in his actions, or ought to exist in his actions, and if that is not the case, it lends him other ideas corresponding to the actions (however perverse) that he does perform" (1971: p. 158).

The flexibility of 'lending other ideas' is what leads us to think of the articulation between the tractor and the trailer as a hinge which can be unhooked and recombined. Thanks to this flexibility we can consider Palestinian stoicism, performed in the linguistic practice of alternating between Hebrew loanwords and Arabic equivalents, to be a negotiation of both Israeli securitism and Palestinian nationalism, as explained above. *Articulation* is Stuart Hall's resolution of the connection between agency and structure recognisable to all social scientists. *Articulation* overcomes problems of determinism but does not dissolve relations of power, and allows for reconfigurations, but nevertheless provides a linkage between various, analytically convenient, distinctions of practices which can be thought of together (Hall 1980: p. 65).

In sum:

1. *Ideologies* are systems of ideas that represent subjects' understanding of their relation to the conditions in which they live, the subjects' actions are performances of this understanding, and therefore ideologies are materially evident in the actions;
2. *Articulation* is a conception of the flexible link between ideologies' pull and the various interconnected social practices of subjects;
3. *Language practices* are a particular type of social practice or action identified by empirical documentation of speech using ethnographic methods that also record participants' understandings of interactions, and that moreover discern patterns in and aspects of the practices that are not readily accessible in ordinary experience.

Linguistic anthropology and ethnography can make good on the promise, hinted at by the second meaning of *articulation* in the sense of voicing (Grossberg 1986), to introduce language into the 'materiality of ideology'. The case of the Israeli Hebrew loanword *menahēl* 'boss' in Palestinian Arabic provides material for an analysis that cannot but incorporate the ideologies that represent relations with the conditions of life in the Palestinian-Israeli context. These conditions include economic precarity

for Palestinian day-migrant workers and military control over access and movement in the Occupied Palestinian Territories, combined with limited autonomy in the form of Palestinian institutions. Palestinians understand their relations to these conditions with combinations of stoicism and nationalism in negotiation with Israeli securitism and economic neo-liberalism. It might seem too obvious to apply this explanatory model to *menahēl*, in its semantic field of power relations, but it's a place to start, not least so as not to annoy 'the boss'.

Acknowledgment

References

Agha, A. (2005). Voice, footing, enregisterment. *Journal of linguistic anthropology* 15 (1), 38–59. https://doi.org/10.1525/jlin.2005.15.1.38

Alenat, S. (2010). *Working for Survival: Labour Conditions of Palestinians Working in Settlements*. Tel Aviv: Kav LaOved.

Althusser, L. (1971). Ideology and ideological state apparatuses (notes towards an investigation). In *Lenin and Philosophy and other essays*, tr. by B. Brewster (pp. 121–173). London: NLB.

Barbe, K. (1995). *Irony in context*. Amsterdam: John Benjamins Publishing. https://doi.org/10.1075/pbns.34

Beinin, J. (2001). *Workers and peasants in the modern Middle East*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511612800

Bucholtz, M. & Hall, K. (2005). Identity and interaction: a sociocultural linguistic approach. *Discourse Studies* 7(4–5), 585–614. https://doi.org/10.1177/1461445605054407

Butler, J. (1997). *Excitable speech: A politics of the performative*. Oxford: Routledge.

D'Anna, L. (2018). *Italian loanwords in Libyan spoken dialect*. In this volume.

Du Bois, J. W. (2007). The stance triangle. In Englebretson (ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction* (pp. 139–182). Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.164.07du

Eagleton, T. (1981). Carnival and Comedy: Bakhtin and Brecht. In *Walter Benjamin, or, Towards a Revolutionary Criticism*. London: Verso

Farsakh, L. (2005). *Palestinian labour migration to Israel: Labour, land and occupation*. Oxford: Routledge. https://doi.org/10.4324/9780203431658

Fishelson, G. (1992). *Forming (1970–87) and Destroying (1988–1990) a Common Labor Market*. Tel Aviv: Tel-Aviv University.

Grossberg, L. (1986). On Postmodernism and Articulation: An Interview with Stuart Hall. *Journal of Communication Inquiry* 10(2), 45–60. https://doi.org/10.1177/019685998601000204

Hall, S. (1980). Cultural studies: Two paradigms. *Media, Culture and Society* 2(1), 57–72. https://doi.org/10.1177/016344378000200106

Hall, S. (1996). Race, articulation, and societies structured in dominance. In H. Baker et al. (eds.) *Black British Cultural Studies: A Reader* (pp. 16–60). Chicago & London: University of Chicago Press.

Hasan, R. (2004). Analysing discursive variation. In L. Young & C. Harrison (eds) *Systemic functional linguistics and critical discourse analysis: studies in social change*. London & New York: Continuum.

Hawker, N. (2013). *Palestinian-Israeli contact and linguistic practices*. Abingdon & New York: Routledge.

Henkin, R. (2009). The cognate curse in Negev Arabic: from playful punning to coexistence conflicts. *Israel Studies in Language and Society* 2(2), 169–206.

Hodge, B., & Mansfield, A. (1985). Nothing left to laugh at …': humour as a tactic of resistance. In *Language and the Nuclear Arms Debate: Nukespeak Today* (pp. 197–211). London: Frances Pinter.

Holes, C., & Abu Athera, S. S. (2009). *Ya Kundalizza Rays!: Poetry and Politics in Contemporary Bedouin Society*. Ithaca & Reading: Ithaca Press

Jacobsen, L. B. (2004). *Educated Housewives: Living Conditions among Palestinian Refugee Women*. Norway: FAFO.

Jaffé, A. (2009). *Stance: sociolinguistic perspectives*. USA: OUP. https://doi.org/10.1093/acprof:oso/9780195331646.001.0001

Kroskrity, P. V. (ed.) (2000). *Regimes of language: ideologies, polities, and identities*. Oxford: James Currey Publishers.

Levy, Y. (ed.) (1995). *Oxford English-Hebrew/Hebrew-English Dictionary*. Jerusalem: Kernerman and Lonnie Kahn Publishing.

Palestinian Central Bureau of Statistics (2015). al-iḥṣāʾ al-falasṭīnī yuʿlan al-natāʾiǧ al-ʾasāsīyya li-masaḥ al-quwa al-ʿāmila li-r-rubʿ al-ʾawwal 2015 [Palestinian statistics reveal the basic results of a survey of the labour force for the first quarter of 2015], retrieved on 16/12/2015 from http://www.pcbs.gov.ps

Pappé, I. (2006). *The ethnic cleansing of Palestine*. Oxford: Oneworld.

Rampton, B. (2009). Dell Hymes's visions of enquiry. *Text & Talk-An Interdisciplinary Journal of Language, Discourse & Communication Studies* 29(3), 359–369.

Roy, S. (1999). De-development revisited: Palestinian economy and society since Oslo. *Journal of Palestine Studies* 28(3), 64–82. https://doi.org/10.2307/2538308

Semyonov, M., & Lewin-Epstein, N. (1987). *Hewers of wood and drawers of water: Noncitizen Arabs in the Israeli labor market*. New York: Ithaca Press.

Silverstein, M. (2000). Whorfianism and the linguistic imagination of nationality. In P. Kroskrity (ed.) *Regimes of language: Ideologies, polities, and identities* (pp. 85–138). Oxford: James Currey Publishers..

Turner, M., & Shweiki, O. (eds.). (2014). *Decolonizing Palestinian Political Economy: De-development and Beyond*. Palgrave Macmillan.

Contact-induced change
from a speakers' perspective

A study of language attitudes in Siwa

Valentina Serreli
University of Bayreuth

The article presents the speakers' perception of contact-induced linguistic change in the Egyptian oasis of Siwa, based on data collected during the authors' doctoral research (Serreli 2016). The research explored language attitudes and ideologies in Siwa with a qualitative approach built on sociolinguistic and linguistic anthropological theories. Linguistic change is presented by speakers as a generational variation; it is attributed to the increased contact between the Siwi and Arabic languages that followed the wider socioeconomic change in the community in recent decades. Moreover, Siwi speakers hold a variety of attitudes towards linguistic change, appreciating phenomena perceived as adjustments to the current times, while criticizing those perceived as a betrayal or corruption of their native language.

Keywords: Siwa Oasis, Berber, Arabic, minority languages, language contact, language attitudes, language change, Egypt

1. Introduction

Acknowledging the significance of folk accounts represents a reversal of scholarly assumptions in both linguistics and anthropology (Kroskrity 2004: pp. 498–9; Gal 2006).[1] In sociolinguistics, the study of language attitudes was formalized after Hoenigswald's (1966) claim that the folk view is worthy of study in its own right, and called it "folk linguistics". The discipline then developed as the study of the speakers' perspective on linguistic facts – both their comments on and reactions to linguistic behavior and the mechanisms underlying those comments and reactions

---

1. Special thanks to Dr. Valentina Schiattarella for the fruitful exchange of ideas and to Dr. Stefano Manfredi for his meticulous revision.

(Niedzelsky & Preston 2000). Sociolinguists study attitudinal variation across social groups and categories and the way in which the social meaning attached to linguistic variables can help to explain patterns of use and trajectories of change (Garret 2001; Milroy & Preston 1999). Linguistic anthropologists, on the other hand, drawing on Silverstein's (1979) work, study language ideologies, that is, "cultural conceptions about language, its nature, structure and use, and about the place of communicative behavior in social life" (Gal 2006: p. 179), as a way to better understand the culture where they occur.

The Arabic-speaking world is a fertile field for folk linguistic research (Walters 2006). For example, taking into consideration the insider perspective helps explain the survival of the longstanding Classical Arabic language ideology and the validity of a two-poles model to describe the Arabic sociolinguistic context (e.g., Haeri 2003; Suleiman 2013; Sayahi 2014). Analyzing the different kinds of prestige enjoyed by different varieties of Arabic and the relationship of power between them helps explain dynamics of language use, variation, and change (e.g., Ibrahim 1986).

Studying speakers' attitudes and state ideologies concerning languages is helpful in understanding the sociolinguistic dynamics within multilingual communities, such as those characterized by the coexistence of (and contact between) Arabic and Berber languages in North Africa. In Morocco and Algeria, the countries with the highest percentages of Berber population, Berber varieties enjoyed neither prestige among speakers nor any official status.[2] Activists' decades-long struggles led to the official recognition of Berber and to a partial change in lay speakers' attitudes, which became more positive (Kossmann 2013: p. 30), even if this did not always lead to a behavioral change (Ennaji & Sadiqi 2008: p. 52; Sayahi 2014: p. 18). The Berber community of Siwa, a small minority of the Egyptian population, has historically received little attention from the Egyptian government. Together with the other minority languages of Egypt (e.g., Nubian and Beja), Siwi lacks any official status so far. Moreover, the political claims made by the international Berber movement, which recently reached Siwa, are rejected by the local community (Serreli 2017).

The research conducted in Siwa was aimed at depicting the current sociolinguistic situation in the oasis, according to both the insiders' perspective and the researcher's observation. The research adopted a qualitative approach in order to provide an in-depth understanding of the phenomena under study and to account for individual nuances and details. Data was collected during intensive fieldwork in the oasis, between 2013 and 2015, through both the direct elicitation of information ("direct approach") and the observation of practices in context ("societal treatment approach," Garret 2006, or "anthropological-cultural approach," Preston

---

2. In Morocco, Berber has been a co-official language since 2011; in Algeria, it became a national and official language in the revised constitution of March 2016.

2011). From one side, people's overt comments about their own linguistic identity and the place and values of Siwi and Arabic varieties in the community were collected through informal conversations and formal interviews.[3] From the other side, the technique of participant observation proved very fruitful as it allowed the researcher to participate in the community's daily life and establish a relationship of trust with the informants that provided the tools for a deeper understanding of both their statements and their behavior.

This article will focus on the perceived contact-induced change in Siwa. Specifically, the article presents speakers' perception of and attitudes towards the variation and change in Siwi, after a brief overview of the sociolinguistic background of Siwa oasis.

2.   Brief overview of Siwa

2.1   General background of the oasis

Situated in the Egyptian Western Desert 50 kilometers away from the Libyan border, the oases of Siwa and El-Gara represent together a unique Berber-speaking enclave in Egypt. The population of 28,329 people includes the Siwan Berbers, regarded by themselves and by others as the "true locals"; an Arab Bedouin tribe, affiliated to the Awlād ʿAli of the Northwestern Coast, which settled in Siwa in the early 20th century (Stein & Rusch 1978: p. 118; Bliss 1984: p. 57); and some thousands of Egyptian outsiders who moved to Siwa for employment purposes from different regions of the country.[4] No data about the exact demographic weight of the groups is available. The local conception of the oasis' population as constituted by three groups, referred to as *siwiyyīn* (Siwans), *badu* (Bedouins), and *maṣriyyīn* (Egyptians) in Arabic, is deeply rooted and automatically acquired.[5] Even outsiders

---

3.   Sixty-one recorded interviews were conducted, both unstructured and semi-structured, of variable length (5 minutes to 2 hours), because they were adapted to each interviewee. Interviewees are male (3/4) and female (1/4), aged between 9 and 70 years. Educational background also varies from uneducated to PhDs. Most of the interviewees are Siwi speakers (3/4); the rest are speakers of Bedouin or Egyptian Arabic.

4.   Data obtained informally from officers of the Governorate of Marsa Matruh in 2015. The 2006 census recorded a population of 21,693 inhabitants (Arab Republic of Egypt – Central Agency for Public Mobilization and Statistics: http://www.msrintranet.capmas.gov.eg/pls/indcs/cnsest_a_sex_fay?LANG=1&lname=FREE&YY=2006&cod=06&gv=33. Retrieved on 11/09/2011).

5.   Hereafter "Siwan(s)" is used to designate the Berber part of the oasis' population, "Bedouin(s)" to designate the Arab segment, and "Egyptian(s)" to designate other Egyptians who do not belong to Arab Bedouin tribes. For a list of Siwi ethnonyms in use in the oasis see Souag (2013: p. 16).

living in Siwa for short periods may start identifying people with one of these labels, regardless of the fact that all of them are actually Egyptian citizens.[6]

The traditional tribal system remains important in the social organization of the community; the eleven tribes who live in the oases, ten Siwan and one Bedouin, comply with the customary law (*'urf*) when dealing with internal issues that do not involve outsiders. However, Siwans and Bedouins remain distinguished by specific social and cultural practices, such as language and clothing. Also, the intermarriage rates are low, and the groups' members overtly express their pride in their own language and culture, this last expressed in terms of *'adāt w taqalīd* (customs and traditions). "Egyptians" do not constitute a tribe, as they moved to Siwa individually or with their nuclear family; they distinguish themselves from (and are distinguished by) both Siwans and Bedouins by a number of practices, such as maintain their native Arabic dialect and their own clothing choices. However, the number of Egyptians residing in Siwa continues to increase, and while Bedouins mostly maintain the ancient custom of marrying parallel patrilineal cousins (*mask bint il-'amm*), intermarriage between "Egyptians" and Siwans becomes more frequent, thus opening new scenarios of cultural and linguistic contact between the two groups.

2.2   Sociolinguistic overview

Travelers who have visited Siwa since the end of the 19th century (e.g., Jennings-Bramly 1897; Wākid 1949; Fakhry 1973), as well as contemporary scholars (Battesti 2006; Souag 2013; Schiattarella 2015; Serreli 2016), describe Siwi as the community's intra-group code of communication. They point out that in the first half of the 20th century the population of the oasis was largely monolingual (e.g., Hohler 1900; Belgrave 1923; Cline 1928; Simpson 1929; Fakhry 1973).

Beginning in the second half of the 20th century, Arabic spread faster and more widely among the oasis' population, triggering a process of transition towards Siwi/Arabic bilingualism. The factors underlying the diffusion of Arabic among Siwans are many and varied: (1) schooling, (2) regular recruitment into the Egyptian army after the 1952 Revolution (Fakhry 1973: p. 36; Cole & Altorky 1998: p. 72; Ellis 2012: pp. 61–62), (3) diffusion of television in Siwans' houses, starting in the 1980s, (4) opening of a 300-km-long asphalt road connecting Siwa to Marsa Matruh

---

6.   The following short anecdote can convey what is meant here. After spontaneously labelling non-Siwans as "Egyptians" in opposition to Siwans, a highly-educated interviewee felt the need to specify how he was using the terms and, without being asked, he commented: "In nationality we are all Egyptian, of course, but I say 'Egyptian' or 'Siwan' based on the language difference" (Int.-10).

(1980s), (5) presence of an increasing number of nonlocal Egyptian workers and public employees, (6) increasing intermarriage rates, (7) increasing number of Siwans attending university in an Egyptian city, where they acquire fluency in the local Arabic dialect, and (8) development of a touristic sector (1980s – 1990s). These factors are by no means specific to Siwa; they are also among the causes of the transition towards bilingualism in a great number of Berber-speaking communities across North Africa (e.g., Chaker 2008: p. 14; Ennaji 2005: p. 71; Kossmann 2013: p. 38).

Today, Arabic is widespread and almost all the population, with the exception of some preschool aged children and elderly women, speaks some form of Arabic in addition to Siwi. As in other Berber-speaking regions (Kossmann 2013: p. 36), Arabic proficiency varies greatly, from the fluency of mixed families' offspring or highly educated people to the passive knowledge of elderly women who learned Arabic from television, but very rarely – or never – have the opportunity to speak it. Gender is a relevant variable for Arabic proficiency principally with elderly and adult speakers: elderly women are less fluent than their male peers, but gender differences progressively diminish among younger generations because girls nowadays come into contact with Arabic as much as boys, thanks to schooling, exposure to television broadcasting, increasing intermarriage rates, and contact with Arabic-speaking neighbors.

Despite Arabization, Siwi continues to be spoken within the community, it is transmitted to offspring and preferred in intra-group interactions by the majority. The field research revealed the emergence of different practices among a minority of young and educated Siwans, who choose to transmit Arabic to their offspring because of its wider diffusion and its perceived greater appropriateness for modern life. However, the intrusion of Arabic into the intra-group domain provokes negative reactions among most members of the community, whose tacit norms still prescribe the use of Siwi among members as a way to signal belonging, equality, and mutual solidarity.

2.3   Siwi and Arabic in contact

Siwi is the easternmost Berber language and it belongs to the Eastern Berber subgroup (Souag 2013: pp. 17–18). According to Kossmann (2013: p. 417), Siwa and Ghomara Berber (in northwestern Morocco) display the highest percentages of lexical borrowing among the Berber languages. According to Souag (2013), most of the Siwi lexicon consists of Arabic loanwords, with different strata of borrowing. Laoust (1932: p. 35) suggests that Siwi was principally influenced by the Bedouin dialect currently spoken in and around Siwa, a Sulaymi Bedouin dialect whose

features include the /g/ reflex of *qāf*, the *gahawa*-syndrome, and the preservation of the feminine plural agreement marking (Souag 2009). On the other hand, Souag (2009) points out that certain phonological features, such as the /q/ reflex of Arabic *qāf* and the final *ʾimāla* in a number of loanwords, seem to link the Arabic element in Siwi to the dialects of other oases. He claims that this element reflects an extinct Arabic dialect spoken in Siwa, whose existence would be consistent with the Arab geographer Al-Idrisi's account of the presence of an Arabic-speaking community in Siwa in the 12th century. Therefore, Souag (2013: pp. 33–34) argues that "neither Classical Arabic alone nor any combination of modern Arabic dialects is individually adequate to account for the Arabic element of Siwi" and he emphasizes that, although the modern contact has caused an Arabic influence in Siwi at the lexical level, "its effects should not be exaggerated."

Schiattarella (2015) argues that, despite being used by most Siwans and being recognized as their mother tongue, Siwi should be considered endangered. She attributes this endangerment to the contact with Arabic, specifically to the growing intermarriage rates and the displacement of many young speakers to Arabic-speaking cities.

Before these last decades of widespread Arabization, the target Arabic variety for Siwans was a Sulaymi Bedouin dialect, spoken by the ʾAwlād ʿAli tribesmen with whom they conducted trade. Now their target Arabic variety is Egyptian Arabic, which they equate with Cairene Arabic. It was observed during the fieldwork that the speech of most of the older men retains Bedouin traits while the speech of most of the younger Siwans displays Egyptian Arabic traits (e.g., /g/ for *qāf* and /ǧ/ for *ǧīm versus* /ʾ/ and /g/, to mention the most significant distinguishing traits according to the speakers themselves). Siwans practice and the targeted Egyptian Arabic do not perfectly coincide, because the speakers sometimes fail to learn some Egyptian Arabic features or carry over features of their native Siwi language. This phenomenon, referred to as "substratum interference" (Versteegh 2001: p. 471) or "shift-induced interference" (Thomason 2001: p. 75), remains unstudied in Siwa so far, and the Arabic spoken by Siwans deserves further investigation. However, this article covers only the perceived contact-induced variation in Siwi.

3.   Perceived contact-induced change in Siwi

3.1   Introductory considerations

The data reported hereafter was discussed in Serreli (2016, Chapter 4). It concerns the speakers' perception of and attitudes towards the ongoing variation and change

in Siwi and does not describe actual phenomena of variation and change from a specialist's point of view.

The discussion draws on a number of interview extracts, which are reported and commented on. For every extract, the interviewee's statements are preceded by his/her first initial, and also the interviewer's remarks are indicated with her first initial (V.). Overall, twenty interviewees are quoted. Table 1 provides details about the interviewees' sex, age, first language, degree of education and occupation at the time of the interview. The last column shows the code of the extracts as reported in the text.

Table 1. Profiles of interviewees

	Name	Sex	Age	First language	Education	Occupation	Extract nr.
Int.-1	F.	f	late 50s	Sw	uneducated	housewife	(1) (2)
Int.-2	A.	m	mid-40s	Sw	mid-educated	curator	(3)
Int.-3	A.	f	late 40s	BdAr/Sw	low educated	housewife	(4)
Int.-4	K.	m	late teens	Sw	low educated	workman	(5) (12)
Int.-5	M.	f	late 20s	EgAr/Sw	high educated	government employee/cultural	(6) (13)
Int.-6	M.	m	early 30s	Sw	high educated	teacher	(7)
Int.-7	O.	m	mid-50s	Sw	high educated	chieftain	(8) (11)
Int.-8	A.	m	late 20s	Sw	high educated	government employee/municipal	(9) (24)
Int.-9	A.	m	mid-30s	Sw	mid-educated	handicraft shop owner	(10)
Int.-10	M.	m	mid-30s	Sw	high educated	receptionist	(14) (17) (30)
Int.-11	S.	m	mid-40s	Sw	high educated	teacher	(15)
Int.-12	O.	m	early 20s	Sw	low educated	military service	(16)
Int.-13	M.	m	mid-30s	Sw	high educated	government employee/municipal	(18) (21) (31)
Int.-14	G.	m	mid-50s	EgAr	mid-educated	hotel owner	(19) (27)
Int.-15	M.	m	mid-20s	EgAr/Sw	high educated	mechanical	(20)
Int.-16	A.	f	40s	Sw	low educated	housewife	(22)
Int.-17	M.	m	early 40s	Sw	mid-educated	farmer	(23) (25)
Int.-18	H.	m	mid-60s	Sw	high educated	government employee/municipal	(26)
Int.-19	A.	m	mid-60s	Sw	high educated	teacher	(28) (32)
Int.-20	Y.	m	late 20s	Sw	mid-educated	hotel owner	(29)

With the exception of the interview with A. (Int. 19), which was conducted in both English and Egyptian Arabic, all interviews quoted were conducted in Egyptian Arabic. However, because of space constraints, the extracts are reported in English translation; only for the most significant expressions are the original Egyptian Arabic words given.

3.2   Presentation of data

3.2.1   *General patterns of Siwi variation*
The oasis' dwellers consider Siwi as a homogeneous or unique language (*luġa waḥda*):

(1)   F. (Int.-1): Within the Arabic language there is Standard Arabic and so on. As for Siwi, no, it is a unique language (*hiyya luġa waḥda*).

The speech of El-Gara Oasis, which is the furthest and most isolated of the Siwi-speaking villages, represents the only exception to such linguistic unity, according to many informants. It is considered to be a form of Siwi with some degree of difference (a different "accent" – *aksent*) from that spoken in Siwa, but there seems to be disagreement about the extent of the difference, as young interviewees aged around 20 years (see (5) and (6)) perceive a bigger difference than older informants do (see (2) and (3)). El-Gara speech is perceived by some interviewees from Siwa as "faster" (*sarīʿ giddan; bi-surʿa*), "weird" (*ġarīb*), and "funny" (*muḍḥak*).

(2)   F. (Int.-1): In Siwa all [villages] speak the same speech (*il-kalām wāḥid*); there is no difference. There is an area called El-Gara. This El-Gara …differs a little bit from the Siwi …but in all Siwa we speak Siwi.

(3)   A. (Int.-2): El-Gara's speech (*il-aksent)* is a little bit different … the sound *(il-fonetiks)* differs a little bit … you understand that one is from El-Gara, it is the same word (*nafs il-kilma*) … the pronunciation (*il-fonetiks*) differs a little bit.

(4)   A. (Int.-3): Only El-Gara … El-Gara is on its own. It is Siwi but different […] their speech is funny (*muḍḥak kalāmhum*).

(5)   V.: Is the Siwi spoken in El-Gara exactly like yours, or different?
       K. (Int.-4): No, it is different (*muxtalif*) … I mean, they speak Siwi but very fast (*sarīʿ giddan*) […] to the extent that you won't understand what he says! … his speech is very weird (*kalāmu ġarīb giddan*)!

(6)   M. (Int.-5): They speak differently (*biykallimu muxtalif*) …I don't know what they say, I didn't learn it (*miš baḥfaẓ*), but I understand if they say two or three words … but they speak very fast (*biykallimu bi-surʿa*).

An interviewee from El-Gara explains how variation is related to the different pace of change between Siwa and El-Gara. He states that the two speeches were originally undistinguishable, as were their speakers' lifestyles, until the speech of Siwa started to change because of the growing contact with Arabic due to the mixture (*ʾixtilāṭ*) with outsiders. Since El-Gara has not yet undergo such changes, their speech remained more conservative.

(7) M. (Int.-6): Now with the contact (*al-ʾixtilāṭ*), the fact that Siwans mix (*ixtalaṭu*) with outsiders more and there started to be change (*taḥrīf*), there is a difference between our language and the language of Siwa. There is an evident difference (*fī ʾixtilāf bāyin*) now because we are still living a simple life (*ʿayišīn al-fiṭra*) until now – that is, until now we are still living basically in a simple way, while here in Siwa, after this contact, they started to introduce words (*yidaxxalu kalimāt*) like this … there is a difference, now there is.

V.: Do you mean that there are, here among Siwans – here in Siwa – people that introduce Arabic words (*kalimāt ʿarabiyya*)?

M.: Yes, they introduce words, while among us, there in El-Gara no. The original (*il-ʾasās*), I mean the original [speech] is still there until now … as for the elderly, it's normal, for the elderly it's the same among us and in Siwa; for my generation (*il-gēl bitāʿi*) it's different! … When I finished the literacy classes and I came here for preparatory school I felt that there was difference (*fī farq*) … This goes back to the year 1992…

V.: Did they feel that your speech was different from theirs?

M.: Yes, they felt that it was different (*muxtalif*) and I felt that there was a difference (*ʾixtilāf*) … they considered mine broken (*mukassar*) and I looked at theirs as broken (*mukassar*).

This idea that linguistic change is directly related to the contact with Arabic speakers and the resulting lifestyle change is widespread among Siwans. Several of the interviewees (see (8), (9) and (10)) suggest a causal relationship between the opening to the outside (*ʾinfitāḥ*) and the innovations (*ʾistiḥdāṯ*) in oasis life and the changes (*taġyīr; at-taġayyurāt*) and innovation (*al-mustaḥdasāt*) in the language itself, which resulted in the adoption (lit. 'entrance' *daxalit; ḥatxošš; daxīla ʿala*) of Arabic loanwords in Siwi. However, besides this alleged natural adaptation to changing times, A. (10) also points out that adopting Egyptian Arabic – or, as he puts it, imitating (*it-taqlīd*) Egyptians' speech – could be seen as an attempt to appear "civilized" (*mutaḥaḍḍar*) because of the association of Egyptianness with civilization and urbanness.

(8) O. (Int.-7): [Siwans] maintain the language as a community, but with individuals it has changed with the opening (*ʾinfitāḥ*) of the oasis – as I said, the opening of the oasis and the innovation (*ʾistiḥdāṯ*) of the things that circulate.

(9) A. (Int.-8): There will be a lot of change (*ḥayibʾa fi taġyīr giddan*) in the language, because Arabic entered (*daxalit*) more than it used to. […] There will be more Arabic [in it] than this, but the language itself won't be lost, won't be lost easily. I mean, with time there will be words that enter it (*ḥatxošš fīha*) maybe, specific words entering in it.

(10) (Int.-9): The language is still there but with changes (*at-taġayyurāt*) or inno-
vations (*al-mustaḥdasāt*). I mean, now when I speak at home I use more [than
before] many words from the Arabic language borrowed into (*daxīla ʿala*) the
Siwi language.

V.: Are there many Arabic loanwords in Siwi (*kalimāt ʿarabiyya daxīla ʿala
s-sīwiyya*)?

A.: A lot! There are a lot! Because we watch television a lot, speak on the phone
a lot, read books, go to school, sit with Egyptian people (*nās maṣriyyīn*) a lot …
and there is something like imitation (*taqlīd*) or emulation of the other – I
mean, I want to show that I am civilized (*mutaḥaḍḍar*).

According to O. (11) and many other interviewees, the new lifestyle brought by
the opening to the outside implies an acceleration of the pace of life, so that adults
are "busy with life" and "don't have time to sit with children". In the next extract
(11), O. refers to the old habit according to which, in the afternoons, men sat in
the shade of a canopy made of palm branches to weave palm-leaf baskets and chat:
when children sat with them and listened to their stories, they acquired both the
language and the traditional Siwan *savoir-faire*. Nowadays, this and other practices
favoring the intergenerational transmission of Siwi, like women's storytelling, are
almost lost.

(11) O. (Int.-7): With all the opening (*ʾinfitāḥ*) of Siwa, the language is affected
(*bitiṯʾassar*) … The young people used to come and sit with [the elderly]: here
they picked up the language from them (*biyāxud minnu l-luġa*) …. More
recently, all the people have become busy with life, with the world. I mean all
of us work, each one of us is busy and doesn't even find time to sit with his
children, and our sons pick up any language from anybody (*ʾibnu yišūf ʾayyi
luġa ʿand ʾayyi ḥadd*).

Language change seems to be perceived as a generational variation and, more
specifically, as a corruption of Siwi, as expressed in extracts (12) and (13) below.
Moreover, the interviewees (see (12) and (14)) identify a generational difference in
speech rhythm: slower (*ti'īl; bi-r-rāḥa*) among the elderly and faster (*ʾaxfaf, xafīf;
sariʿa šwayya*) among youth.

(12) K. (Int.-4): There is a difference, you know, between someone who is eighty
years old and a young person who's twenty years. I mean, the young one grew
up with Egyptian and Siwi, even his speech is faster, fast (*ʾaxfaf, xafīf*). But the
old man's speech is thick (*ti'īl*) … his Siwi speech is thick (*ti'īl*), slow (*bi-r-rāḥa*).

(13)   M. (Int.-5): I go to listen to the elderly to know what's right and what's wrong (*illi ṣaḥḥ w illi ġalaṭ*) … I mean, now I don't know which is the correct language (*il-luġa ṣ-ṣaḥḥ*), so I ask the elderly. Nobody knows except maybe ninety- or eighty-year-old people, you feel that someone like that might tell you the correct thing, but people at the age of my mother don't know everything that's correct … fifty-two years old, she doesn't know.

(14)   M. (Int.-10): Nowadays … they speak the dialect a little bit fast (*sariʿa šwayya*).

The perceived variation in Siwi can be stated in terms of (1) the loss of Siwi lexicon related to the traditional life, (2) the integration of Arabic loanwords into Siwi, and (3) cases of codemixing and use of "broken Siwi".

3.2.2   *Loss of Siwi lexicon related to the traditional lifestyle*

Siwi speakers consider Siwi as the perfect means to express traditional Siwan life in all its aspects. With the abandoning of traditional daily activities and jobs, the lexicon referring to them is no longer needed nor used and, therefore, unknown to younger people. The interviewees quoted in (15), (16), (17), and (18) point out that any young speaker who never saw old instruments, who was neither involved in nor passively witnessed traditional activities (*ma-taʿwuttš ʿalēha wa la šafha*), did not learn the Siwi lexicon that referred to it (*musṭalaḥāt ʾadīma; kalimāt; musṭalaḥāt*), which has consequently been lost or forgotten (*yanqarid; nasyat*).

(15)   S. (Int.-11): Now some Siwi terminology is starting to disappear (*yanqarid*) […] Some expressions enter (*daxal musṭalaḥāt*) from Arabic as substitutes (*badayil*) because of their frequent use and there is old terminology (*musṭalaḥāt ʾadīma*) that I actually know, but the generation after mine doesn't know. My children are like this!

(16)   O. (Int.-12): The new generation, Valentina, there are many words (*kalimāt katīra*) that they won't know …things like most of the tools used in the gardens … the new generation, as they don't go to the gardens … there are people that never go to the gardens. … If one's father is a government employee, he comes from school, from private lesson to school, from private lesson to the house … I mean you might ask anyone from the new generation and they would tell you that they don't know, but if [you ask] an older person he does [know]!

(17)   M. (Int.-10): There is a great alteration (*taʾsīr kabīr*)! … Some children have forgotten many Siwi words (*biyinsu kalimāt katīra sīwiyya*) … and also young people now! They forgot even… I mean there are things, if we go to the Siwan House[7] there are many things from the [cultural] heritage that today's youth

---

7.   Ethnographic museum situated in the center of Siwa that displays traditional objects and handicraft.

don't know, nor do they know what such things are used for! … They have neither got used to nor seen them (*ma-taʕwuttš ʕalēha wa la šafha*). By the way, many customs and traditions have become extinct (*yinqaradat*).

(18)   M. (Int.-13): There is some terminology (*muṣṭalaḥāt*) that was used before and now is not …. For example, the elderly people remember things that aren't used now, that's why you feel that they speak … That's it, an item that is not used is forgotten (*ḥāga ma-btistaxdamš fa nasyat*) … You ask me if the elderly differ from young people, and I told you no, but there are some simple terms or small things (*muṣṭalaḥāt basīṭa aw ḥāgāt basīṭa*) […] that were used in the past but now they are not; nobody uses them and young people won't know them.

This kind of loss is presented as the natural outcome of the changing times. Nonetheless, some interviewees (see (16) and (17)) display nostalgia for the past and show regret for the loss of lexicon. These interviewees usually locate the innovation within an entire age category rather than attribute it to individual speakers, using the collective nouns "new generation" or "youth". However, O., both in the extract reported above (16) and elsewhere in his interview, emphasizes that it is not only a matter of age, but it also depends on the family's background and, more specifically, on the occupation of one's father and how closely it is related to traditional activities and tools. For O., as well as for other interviewees, the social distribution of this lexical loss corresponds to other phenomena, such as the degree of Arabic fluency. As we will see below, it applies particularly to the integration of Arabic loanwords.

3.2.3   *Integration of Arabic loanwords into Siwi*

Borrowing consists in the integration of elements from a foreign language into one's native language (Thomason & Kaufman 1988: p. 21). We will adopt Kossmann's (2013: p. 89) distinction between "additive borrowing", which fills a lexical gap in the recipient language when new concepts are adopted, and "substitutive borrowing", which "substitutes or creates an alternative to an existing term."

Siwi speakers recognize the existence of borrowing from Arabic into Siwi. In most cases, they refer to additive borrowing and deny the presence of substitutive borrowing. Some interviewees – as in (19), (20) and (21) – account for the integration of Arabic loanwords as a way to fill a lexical gap, overlooking or refusing any relationship to the prestige enjoyed by Arabic that renders its use fashionable (*mōḍa*) among youth. Other interviewees (see (22)) point out that the Siwan youth are fascinated by Egyptian Arabic (as well as by the Egyptian lifestyle, as she stated in other moments of the interview) and are willing to adopt Egyptian Arabic loanwords.

(19) G. (Int.-14): He speaks Siwi and he may introduce … even the words are new, because before there were no televisions or refrigerators or washing machines, either … all these things weren't here!
V.: Is it because of this that new things arrived with an Egyptian name (*bi-ʾism maṣri*)? ?
G.: Yes, with the Egyptian name (*bi-l-ʾism ǝl-maṣri*)!
V.: Is there substitution (*badīl*) also? … The things that were already there …
G.: No, I don't think so.

(20) M. (Int.-15): Many Arabic words entered (*daxalat*) […]
V.: But are the loanwords (*il-kalimāt il-daxīla*) words referring to new things or they are words borrowed just because they're in vogue (*kalimāt daxīla wi xalāṣ ʿašān mōḍa*)?
M.: […] No, just the new things.

(21) M. (Int.- 13): A lot of expressions […] entered (*musṭalaḥāt katīra […] daxalat*). This terminology is … for example, the devices, all the modern devices (*al-ʾaghiza l-mustaḥdasa l-gadīda*), like the computer and the radio (*il-kumbiyūtar w il-radiyo*), all these things entered by the same name … I mean, they kept the same name as they entered with …
V.: And what about the things that were already there? … Is it possible that an Arabic word enters (*tadxul*) as a substitute for (*yibādil*) a Siwi word?
M.: No, it isn't but the modern things (*il-ḥāgāt il-mustaḥdasa*) I told you about, but Siwi … I mean the Siwi words remained (*faḍalit*) Siwi. But the Siwan nowadays, he speaks his language, Siwi, and puts in a little bit (*yḥōṭṭ šwayya*) of Arabic because, as I told you, all modern things (*il-ḥāgāt il-ḥadīsa*) …

(22) A. (Int.-16): There are words now that they pronounce (*biyanṭaʾūha*) in Egyptian Arabic … like '*laptop*', '*tilvizyōn*': these all are Egyptian Arabic words. I mean, the new things they see, they name them (*biyanṭaʾūha*) in Egyptian now, there isn't Siwi and Egyptian, all [the things brought by] progress, they name (*biyanṭaʾu*) them in Egyptian. … They arabize (*yiʿarrabu*) the Siwi language, everything that appears (*tiẓhar*) now, they say it (*biyikallimūha*) in Egyptian … all the new generation, they want to speak Egyptian basically, they want to speak Egyptian!

In fewer cases – as in (23) and (24) – the interviewees refer to the replacement of old Siwi words with Arabic loans. However, in these two extracts it is not very clear whether they are talking about the loss of old Siwi lexicon whose referees are not used anymore and the integration of Arabic terms for newly arrived items, or whether they are pointing to the integration of Arabic loanwords as a replacement of Siwi words for existing items and concepts.

(23) M. (Int.-17): There are names that changed (*taġayyarat*) … they changed, started to be pronounced (*badʾat tunṭaʾ*) in Arabic.

(24) A. (Int.-8): Some Siwi speech (*kalām sīwi*) now is lost (*rāḥ*) and Arabic took its place (*daxal makānu*) … it won't reach (*miš ḥatuwṣal*) the next generation because they didn't hear it at all. Sometimes there are words that my grandmother and my grandfather say that I don't know. Why? Because from the time I was born I got used to hearing it (*taʿwutt ʿalēha basmaʿha*) in Arabic and so that's it, I didn't get used to it. This is the evolution (*taṭawwur*) that is happening in the Siwi language.

Another point commented on by interviewees is the form in which the loanword is taken over – that is, whether it undergoes a change towards a form that fits better into the Siwi morphology or whether it maintains its original form. According to them, both cases are attested and, in the case of adaptation (*taḥrīf*), they indicate the addition of the feminine marker, that is the prefix *t-* and the suffix *-ət* (*t-_-ət*), as one of the possible patterns.

(25) M. (Int.-17): All the modern things that come up don't take a Siwi name anymore. I mean, for example, '*tilifōn*' is like this '*tilifōn*'; '*kitāb*' is '*kitāb*'. Maybe there is a simple variation like '*tallāga*' '*tatlāǧət*'[8] … What happens? The variation (*taḥrīf*) of a part [of the word].

(26) H. (Int.-18): He turns (*yilibʿa*) an Arabic word into Siwi … for example, '*ġurfa*' in Siwi becomes '*tɣəṛfət*'; it is from *ġurfa* (room) in Arabic, but it is Siwi. … *maʿlaʾa* (spoon), '*timaʿlaqt*' this is also new […] it isn't Siwi.

Some interviewees – as in (27) and (28) – present borrowing as unidirectional, that is, as the integration of Arabic words into Siwi:

(27) G. (Int.-14): He speaks Siwi but he may introduce (*mumkin yidaxxal*) a few words from the Egyptian dialect (*ʿāmmiyya maṣri*) … but it doesn't happen that he speaks Egyptian and introduces Siwi words.

(28) V.: Do they speak Siwi with Arabic words or Arabic with Siwi words?
        A. (Int.-19): No, Siwi with Arabic words.

This unidirectionality is explained by the fact that speakers perceive borrowing as an instrument of linguistic adjustment to modern times: as Siwans' lives have changed because of the introduction of new items, the Siwi language also needs to change in order to be adequate to the modern life.

---

8.   In order: 'telephone', 'book', 'refrigerator'.

3.2.4   *Codemixing and "broken Siwi"*

Other alleged outcomes of the contact with Arabic are codemixing and "broken Siwi". Here, "codemixing" refers to the mixing of languages within the utterance, while "broken Siwi" identifies an utterance in Siwi that is judged incorrect by other speakers. Unlike the borrowing phenomena discussed above, these are evaluated quite negatively by the interviewees.

For example, Y. in (29) criticizes children's codemixing, attributed to their young parents' choice to transmit Arabic terms of address, and states that he considers it worse than a shift to Arabic, which he had strongly condemned elsewhere in his interview and in other unrecorded conversations.

(29)   V.: Is it only the words '*bāba*' and '*māma*' that are introduced but the rest of the discourse is Siwi?
Y. (Int.-20): […] Even the discourse is not homogeneous (*muš mutarakkib maʿa baʿḍ*). I mean, how can I say, for example, '*bāba xsix*'[9] … The right things are '*bāba ʿāyiz*'[10] or '*abba xsix*',[11] like this, not '*bāba xsix*'. Now they say '*bāba xsix*' … I don't object that he speaks Arabic, but either he speaks Arabic or he speaks Siwi, not half-and-half.

In other cases, young speakers are held responsible for the corruption of the language due to the influence of Arabic, and blamed for doing it in order to sound more Egyptian. So, in (30) M. denounces a case in which the Siwi word has lost its characterizing elements (interestingly, the same ones that are added to adapt Arabic words into Siwi: *t-* …*-ət*) and taken an arabized form (*tgurgət* = *l-gurg*); in (31) M. accounts for the introduction and adaptation of the Egyptian greeting formula *ʾēh l-ʾaxbār?*/*ʾaxbārak ʾēh?* (how are you?), which became a mixed Siwi/Egyptian expression through the translation into Siwi of the interrogative *ʾēh* (Ar.) = *tanta* (Sw.), the substitution of the personal pronoun *-ak* (Ar.) = *-ənnək* (Sw.), and the adaptation of *ʾaxbār* (Ar.) = *lxḅaṛ* (Sw.).

(30)   M. (Int.-10): For example, the wood here – the one from olives trees – is called '*tgurgət*' or '*tigurga*'. Sometimes young people now say '*l-gurg*' … ah! I mean, this is different … I mean, even the elderly [say] 'What? Why do young people say this word?' 'Why this difference?'.

---

9.   "Dad, I want": mixed Egyptian Arabic and Siwi.

10.   "Dad, I want": Egyptian Arabic.

11.   "Dad, I want": Siwi.

(31) M. (Int.-13): *'tanta əlaxḅaṛ?'* for example, … this is an innovation (*istiḥdās*) … *'tanta'* is Siwi, it is an interrogative particle, a question, I mean (*adāt ʾistifhām, suʾāl yaʿni*); what is *'lxḅaṛ-ənnək'*? It is a mixture between Siwi and the Egyptian dialect (*il-ʿāmmiyya*). … Maybe an elderly person would not say it … this word is a little bit new. I hear Egyptians saying *''ēh l-ʾaxbār?'* and I want to say the same expression (*nafs il-kalām*), because I liked the expression (*il-kalām*). I mean, I want to say the same expression (*nafs il-kalām*) and I say it in my language, in Siwi, and there is a change (*taʾsīr*). … In Siwa there wasn't anything like *'tanta əlaxḅaṛ-ənnək?'* but one hears the Arabic *ʾēh l-ʾaxbār?* and wants to translate it in Siwi.

Sometimes older interviewees denounce young speakers' incorrectness very straightforwardly. In the extract (32) reported below, A. defines young people's alleged corruption of Siwi as a "misuse" of the language: he specifies that he does not condemn the integration of foreign lexical items referring to new things, but instances of what he perceives as incorrect Siwi. He mentions examples of morphological changes and the loss of the lexical richness of Siwi. Showing great dislike and disappointment for these trends, A. attributes them to the "neglect" of Siwi within the community which, according to him, pushes children towards Arabic at too early an age, when they should be learning Siwi before concentrating on Arabic.[12]

(32) (Int.-19): [Siwi] is in danger already! […] The young generation starts to use terminology that didn't exist in Siwi before and they start to misuse the words. […] If they start to use any new language that has to do with computer, that is fine, because they are things that didn't exist before. But let's talk about … they don't differentiate between 'watch' and 'see' … they always use 'watch' for both … and then, I'll tell you another thing: *'i zdəffri'* means 'behind me', *'i zdat-i'* 'in front of me', but they don't say it like this now …
V.: What do they say?
A.: *'zdəffr-ənnaw' 'zdat-ənnaw'* they break the language
…
V.: Why do they do it?
A.: Because now the Siwi is neglected (*muhmil*). They concentrate on Arabic … Because now four-year-old children enter the kindergarten and [teachers] talk to them only in Arabic and the Siwi didn't take enough […] The child didn't learn enough Siwi language […] Nowadays even if the father and mother speak Siwi, the television doesn't speak Siwi. Children are always sitting in front of the television and this is another problem.

---

12. Interview conducted in both Egyptian Arabic and English.

A. (Int.-19) is highly educated, he has traveled a lot, and he has been working in the field of education in Siwa for decades, besides being among the first involved in the sector of ecotourism in the oasis. His statements are representative of the views of an emerging elite of educated and (mostly) young Siwi speakers who worry about the possible loss of Siwi and advocate the preservation of the Siwan cultural heritage in line with discourses of identity, endangerment, and preservation that are already at work globally (Duchêne & Heller 2007).

4.   Conclusion

The longstanding influence of Arabic on Siwi lexicon is recognized by linguists (Souag 2013; Schiattarella 2015). Schiattarella also points out that Siwi should be considered endangered and that new forms of contact with Arabic are likely to trigger further linguistic change in Siwi. Siwi speakers present the variation as a generational change, and mostly attribute it to the changing times and the contact with Arabic that this implies. Overt attitudes towards Arabic influence on Siwi vary according to the alleged motivation underlying it: the attitude is positive if the change is perceived as an adjustment to the changing times, and negative if the change is perceived as a gratuitous "corruption" of Siwi because of the greater prestige enjoyed by Egyptian Arabic.

Siwans praise innovation and lifestyle changes because they represent an improvement in living conditions. Arabization is viewed as one of the outcomes of the opening of the oasis and it is evaluated positively. Therefore, Siwans have a correspondingly positive attitude towards additive borrowing, and they also accept the loss of Siwi lexicon referring to items fallen into disuse, perceived as automatic consequences of social changes.

The success of Arabization and the consequent widespread bilingualism are not (primarily) related to the prestige enjoyed by Egyptian Arabic to the detriment of Siwi, nor are they considered a threat to the status of Siwi as the code of intra-group communication. However, Siwans value Egyptian Arabic as a language of wider communication, indexing educatedness and urbanness, and the prestige it enjoys among youth is growing. Interviewees ascribe to this growing prestige the emergence of such phenomena as substitutive borrowing, codemixing and language change, which are criticized by older Siwi speakers.

The research revealed that, although Siwi speakers value Egyptian Arabic because of its strong association with positive innovation and social change, they continue to perceive the link between Siwi and the community as very strong, even though almost the entire population is bilingual. The criticism of non-necessary lexical intrusions of Arabic into Siwi is part of a general disapproval of the intrusion

of Arabic in domains of language use previously reserved to Siwi. This position is encouraged by discourses of heritage preservation or heritage making, which are gaining ground among an elite of young educated Siwans and counterbalance the opposite trend driven by the prestige that Egyptian Arabic enjoys among other segments of the Siwan youth.

However, we should not assume a perfect correspondence between speakers' statements and their actual behavior. A gap exists between overt declarations and the covert attitudes directing speakers' actual choices and practices, i.e., a speaker's speech could display features that (s)he overtly refuses or denies. Therefore, the data presented here about perceived change does not reveal actual patterns of language variation and change. Its study is important, however, as it offers an insight into Siwans' awareness and conceptualization of their community's linguistic situation and the ideologies they endorse.

References

Battesti, V. (2006). 'Pourquoi j'irais voir d'en haut ce que je connais déjà d'en bas?', Centralités et circulations: comprendre l'usage des espaces dans l'oasis de Siwa. In V. Battesti & N. Puig (eds.), *Égypte/Monde Arabe, Troisième série, 3. Terrains d'Égypte, anthropologies contemporaines* (pp. 139–179).

Belgrave, D. (1923). *Siwa, the oasis of Jupiter Ammon*. London: John Lane.

Bliss, F. (1984). *Kulturwandel in der Oase Siwa*. 2nd ed. Bonn: PAS.

Chaker, S. (2008). Langue. In *Encyclopédie berbère*, 28–29, Kirtēsii–Lutte, Aix-en-Provence: Edisud. Retrieved on 31/05/2014 from http://encyclopedieberbere.revues.org/314.

Cline, W. (1928). Notes on the origin of the people of Siwa and Gara. *Man* 28, 24–25.

Cole, D., & Altorky, S. (1998). *Bedouin, Settlers, and Holiday-Makers. Egypt's Changing Northwest Coast*. Cairo: The American University in Cairo Press.

Duchêne, A. and M. Heller (eds.). (2007). *Discourses of Endangerment: Ideology and Interest in the Defense of Languages*. London: Continuum.

Ellis, M. H. (2012). *Between Empire and Nation: The emergence of Egypt's Libyan Borderland, 1941–1911*. (Unpublished doctoral dissertation). Princeton University, Princeton, N.J.

Ennaji, M. (2005). *Multilingualism, Cultural Identity, and Education in Morocco*. New York: Springer.

Ennaji, M. and F. Sadiqi. (2008). Morocco: Language, Nationalism, and Gender. In A. Simpson (ed.), *Language and National Identity in Africa* (pp. 44–60). Oxford: Oxford University Press.

Fakhry, A. (1973). *Siwa oasis*. Cairo: American University in Cairo Press.

Gal, S. (2006). Linguistic Anthropology. In K. Brown (ed.), *Encyclopedia of Language and Linguistics* (2nd ed.), vol. 7 (pp. 171–185). Oxford: Elsevier.

Garret, P. (2001). Language Attitudes and Sociolinguistics. *Dialogue*, 626–630.

Garret, P. (2006). Language attitudes. In C. Llamas, L. Mulling & P. Stockwell (eds.), *The Routledge Companion to Sociolinguistics* (pp. 116–121). London and New York: Routledge.

Haeri, N. (2003). *Sacred Language, ordinary people: Dilemmas of culture and politics in Egypt*. New York: Palgrave Macmillan.

Hoenigswald, H. M. (1966). A Proposal for the Study of Folk-Linguistics. In W. Bright (ed.), *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference, 1964* (pp. 16–26). The Hague: Mouton.

Hohler, T. B. (1900). *Report on The Oasis of Siva*. London: Waterlow and Sons Limited printers, London Wall.

Ibrahim, M. H. (1986). Standard and Prestige Language: A Problem in Arabic Sociolinguistics. *Anthropological Linguistics* 28(1), 115–126.

Jennings-Bramly, W. (1897). A Journey to Siwa in September and October, 1896. *The Geographical Journal* 10(6), 597–608.

Kossmann, M. (2013). *The Arabic Influence on Northern Berber*. Leiden/Boston: Brill.

Kroskrity, P. (2004). Language Ideologies. In A. Duranti (ed.), *A companion to Linguistic Anthropology* (pp. 496–517). Malden: Blackwell Publishing Ldt.

Laoust, E. (1932). *Siwa I: son parler*. Paris: Ernest Leroux.

Milroy, L. & Preston, D. R. (1999). Attitudes, Perception and Linguistic Features. *Journal of Language and Social Psychology* 18(1), 3–8.

Niedzielsky, N. & Preston, D. R. (2000). *Folk Linguistics*. Berlin: Mouton de Gruyter.

Preston, D. R. (2011). Methods in (applied) folk linguistics. Getting into the minds of the folk. *AILA Review* 24, 15–39.

Sayahi, L. (2014). *Diglossia and Language Contact. Language Variation and Change in North Africa*. Cambridge, UK: Cambridge University Press.

Schiattarella, V. (2015). *Le berbère de Siwa: documentation, syntaxe et sémantique*. (Unpublished doctoral dissertation). École pratique des Hautes Études, Paris.

Serreli, V. (2016). *Society, languages and ideologies in the oasis of Siwa. Listening to people's voices*. (Unpublished doctoral dissertation). University of Sassari, Sassari, and University of Aix/Marseille, IREMAM, Aix-en-Provence.

Serreli, V. (2017). Sentimenti di identità e appartenenza dei berberi egiziani di Siwa. In N. Melis (ed.), *Frontiere, confini e zone di frontiera nella regione MENA. Afriche e Orienti* XIX (2), 53–70.

Silverstein, M. (1979). Language Structure and Linguistic Ideology. In P. Clyne, W. Hanks, & C. Hofbauer (eds.), *The Elements: a Parasession on Linguistic Units and Levels* (pp. 193–247). Chicago: Chicago Linguistic Society.

Simpson, G. E. (1929). *The heart of Libya; the Siwa Oasis, its people, customs and sport*. London: H.F. & G. Witherby.

Souag, L. (2009). Siwa and its significance for Arabic dialectology. *Zeitschrift für arabische Linguistik* 5, 51–75.

Souag, L. (2013). *Berber and Arabic in Siwa (Egypt). A Study in Linguistic contact*, Köln: Rüdiger Köppe Verlag.Stein, L., & Rusch, W. (1978). *Die Oase Siwa. Die Oase Siwa. Unter Berbern und Beduinen der Libyschen Wüste* Leipzig: VEB F.A. Brockhaus Verlag.

Suleiman, Y. (2013). Arabic Folk Linguistics: Between Mother Tongue and Native Language. In J. Owens (ed.), *The Oxford Handbook of Arabic Linguistics* (pp. 264–280). Oxford: Oxford University Press.

Thomason, S. G. (2001). *Language Contact*. Edinburg: Edinburg University Press.

Thomason, S. & Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, Los Angeles, and London: University of California Press.

Versteegh, K. (2001). Linguistic contact between Arabic and other languages. *Arabica* 48, 470–508.

Wākid, ʿA. (1949). *Wāḥat Amūn: baḥṯ šāmil li-wāḥat sīwa*. Cairo: Maṭbaʿat al-Muqtaṭaf.

Walters, K. (2006). Language Attitudes. In K. Versteegh (ed.), *Encyclopedia of Arabic Language and Linguistics* 2 (pp. 650–664). Leiden: Brill.

Index

The present volume provides an overview of current trends in the study of language contact involving Arabic. By drawing on the social factors that have converged to create different contact situations, it explores both contact-induced change in Arabic and language change through contact with Arabic. The volume brings together leading scholars who address a variety of topics related to contact-induced change, the emergence of contact languages, codeswitching, as well as language ideologies in contact situations. It offers insights from different theoretical approaches in connection with research fields such as descriptive and historical linguistics, sociolinguistics, ethnolinguistics, and language acquisition. It provides the general linguistic public with an updated, cutting edge overview and appreciation of themes and problems in Arabic linguistics and sociolinguists alike.

John Benjamins Publishing Company