

Arabic and contact-induced change

Edited by

Christopher Lucas

Stefano Manfredi

Contact and Multilingualism



Series information: ./langsci/seriesinfo/cam-info.tex not found!

ISSN:

Arabic and contact-induced change

Edited by

Christopher Lucas

Stefano Manfredi

Lucas , Christopher & Stefano Manfredi (ed.). 2019. *Arabic and contact-induced change* (Contact and Multilingualism). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/000>

© 2019, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: no digital ISBN

no print ISBNs!

ISSN:

no DOI

ID not assigned!

Cover and concept of design: Ulrike Harbort

Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono,

ScheherazadeRegOT

Typesetting software: \LaTeX

Language Science Press

Unter den Linden 6

10099 Berlin, Germany

langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität



Berlin

Contents

I Contact-induced change in varieties of Arabic

1 Khuzestan Arabic	
Bettina Leitner	3
2 Anatolian Arabic	
Faruk Akkuş	23
3 Cypriot Maronite Arabic	
Mary Ann Walter	47
4 Maghrebi Arabic	
Adam Benkato	65
5 Moroccan Arabic	
Jeffrey Heath	83
6 Maltese	
Christopher Lucas & Slavomír Čéplö	95
7 Nigerian Arabic	
Jonathan Owens	135
8 Arabic pidgins and creoles	
Andrei Avram	159
9 Arabic in the diaspora	
Luca D’Anna	187

Contents

II Language change through contact with Arabic

10 Beja	
Martine Vanhove	207
11 Berber	
Lameen Souag	229
12 Modern South Arabian Languages	
Simone Bettega & Fabio Gasparini	245
13 Northern Domari	
Bruno Herin	265
14 Mediterranean Lingua Franca	
Joanna Nolan	287

III Domains of contact-induced change across Arabic varieties

15 Dialect contact and phonological change	
William M. Cotter	305
16 Contact-induced grammaticalization between Arabic dialects	
Thomas Leddy-Cecere	321
17 Contact and calquing	
Stefano Manfredi	343
18 Contact and the expression of negation	
Christopher Lucas	361
19 Contact and variation in Arabic intonation	
Sam Hellmuth	387
Index	407

Part I

Contact-induced change in varieties of Arabic

Chapter 1

Khuzestan Arabic

Bettina Leitner

University of Vienna

Khuzestan Arabic is an Arabic variety spoken in the southwestern Iranian province of Khuzestan. It has been in contact with (Modern) Persian since the arrival of Arab tribes in the region before the rise of Islam. Persian is the socio-politically dominant language in the modern state of Iran and has influenced the grammar of Khuzestan Arabic on different levels. The present article discusses phenomena of contact-induced change in Khuzestan Arabic and considers their limiting factors.

1 Current state and historical development

1.1 Historical development

Arab settlement in Iran preceded the Arab destruction of the Sasanian empire with the rise of Islam. Various tribes, such as the Banū Tamīm, had settled in Khuzestan prior to the arrival of the Arab Muslim armies (Daniel 1986: 211). In the centuries after the spread of Islam in the region, large groups of nomads from the Ḥanīfa, Tamīm, ʿAbd-al-Qays, and other tribes crossed the Persian Gulf and occupied some of the territories of southwestern Iran (Oberling 1986: 215). The Kaʿb, still an important tribe in the area,¹ settled there at the end of the sixteenth century (Oberling 1986: 216). During the succeeding centuries many more tribes moved from southern Iraq into Khuzestan. This has led to a considerable increase of Arabic speakers in the region, which until 1925 was called Arabistan (see Gazsi 2011: 1020; Gazsi, this volume). Today Khuzestan is one of the 31 provinces of the Islamic Republic of Iran, situated in the southwest, at the border with Iraq.

There has been considerable movement to and from Iraq, to Kuwait, Bahrain, and Syria, and from villages into towns. Many of these migrations were a consequence of the Iran–Iraq war (1980–1988), but some were due to socio-economic

¹Cf. Oberling (1986: 218) for an overview of the Arab tribes in Khuzestan.



Bettina Leitner

reasons. The settlement of Persians in the region over the past decades (Gazsi 2011: 1020) is another important factor in its demographic history. From the early twentieth century on, Khuzestan has attracted international, especially British, interest because of its oil resources.

1.2 Current situation of Arabs in Khuzestan

Information about the exact number of Arabic-speaking people in Iran, and especially in Khuzestan, is hard to find. Estimates in the 1960s of the Arabic-speaking population in Iran ranged from 200,000 to 650,000 (Oberling 1986: 216). Today it is estimated that around 2 to 3 million Arabs live in Khuzestan (Matras & Shabibi 2007: 137; Gazsi 2011: 1020).

Many Arabs and Persians living in Khuzestan work in the sugar cane or oil industries, but few hold white-collar or managerial positions (De Planhol 1986: 55–56). This is one of the reasons why many Arabs in Khuzestan feel strongly disadvantaged in society and politics in comparison to their Persian neighbours.²

2 Language contact in Khuzestan

Currently, the main and most influential language in contact with Khuzestan Arabic is the Western Iranian language Persian. Among the other (partly historically) influential languages in the region the most prominent are English, Turkish/Ottoman (cf. Ingham 2005), and Aramaic (see Procházka, this volume).

Persian and different forms of Arabic share a long history of contact in the region of Khuzestan, implying a long exchange of language material in both directions.

Khuzestan Arabic belongs to the Bedouin-type south Mesopotamian *gəlat*-dialects.³ Therefore, it shows great similarity to Iraqi dialects such as Baṣra Arabic, as well as to other dialects in the Gulf, such as Bedouin Bahraini Arabic – that is, the Arabic spoken by the Sunni Arab population descended from Najd.

²The most common Khuzestan Arabic terms for the Persian people and their language are *ṣaḡam* ‘Persian’ (people and language; lit. ‘non-Arab’), and *al-ḡamāʿa* ‘Persians’ (lit. ‘group of people’). Both are often used pejoratively.

³There is as yet no comprehensive grammar of the dialects of Khuzestan. The main source of information on these dialects is the collection of data made in the 1960s by the arabist and linguist Bruce Ingham (1973; 1976; 2011). The article by Yaron Matras and Maryam Shabibi, “Grammatical borrowing in Khuzistani Arabic” (Matras & Shabibi 2007), is based on Shabibi’s unpublished dissertation “Contact-induced grammatical changes in Khuzestani Arabic” (Shabibi 2006).

1 *Khuzestan Arabic*

The dialects of Khuzestan can be considered “peripheral” dialects of Arabic because they are spoken in a country where Arabic is not the language of the majority population and is not used in education and administration. Therefore, there is practically no influence of Modern Standard Arabic. However, because it shares a long geographically-open border with Iraq, Khuzestan is not isolated from the Arabic-speaking world. Moreover, since around 2000 it has had access to Arabic news, soaps, etc. via satellite TV. Intra-Arabic contact is limited to the linguistically very similar (southern) Iraqi dialects⁴ through, for example, religious visits to Kerbala.

Persian is the only official language in Iran, and the only language used in education, and is sociolinguistically and culturally dominant, especially in the domains of business and administration. Persian consequently enjoys high prestige in society. For Persian speakers, and sometimes also for Khuzestan Arabic speakers, the Khuzestan Arabic varieties have very low prestige and are not associated with the highly prestigious Arabic of the Quran, which is taught at schools. Khuzestan Arabic speakers who acquire Khuzestan Arabic as a first language usually acquire Persian at school. Later, the opportunities for Khuzestan Arabic speakers to use Persian are restricted to certain social settings outside the family, e.g. school, work (employment in a large company would probably require communication in Persian), contact with Persian friends, or through the Persian media.

Accordingly, the command of Persian or the degree of bilingualism among Khuzestan Arabic speakers varies greatly due to such factors as level of education, affiliation, age, gender, and urban or rural environment. The older generation and women have far less access to education and jobs and consequently less contact with people outside the family, which implies less exposure to contact situations and a lower degree of bilingualism. Among some members of the younger generation we may notice a certain intentional reinforcement of Arabic words alongside a resistance to recognizable Persian lexical borrowings plus a preference for the Arabic over the Persian names for the cities in Khuzestan. This of course goes in line with nationalist ideas and the separatist movement taking place in present-day Khuzestan and also shows the impact of intentionality in language contact situations.

In sum, one might find very different degrees of Persian influence among the speakers of Khuzestan Arabic (cf. [Matras & Shabibi 2007](#): 147). For that reason,

⁴Khuzestan Arabic is often differentiated from its neighboring Iraqi dialects by the number of Persian borrowings that are employed ([Gazsi 2011](#): 1020). Although the greatest influence has occurred in lexicon, Persian influence also extends to grammar (see below).

Bettina Leitner

all statements on Persian–Khuzestan Arabic contact phenomena must be seen in relation to the above factors, which are decisive for any speaker’s command of Persian.

3 Contact-induced changes in Khuzestan Arabic

3.1 General remarks⁵

The main aim of the present chapter is to highlight the most striking phenomena and trends in Khuzestan Arabic language change due to contact with Persian.

All phenomena of contact-induced change in Khuzestan Arabic can be considered as transfer of patterns or matter⁶ from the source language (SL) Persian to the recipient language (RL) Khuzestan Arabic under RL agentivity (i.e. borrowing rather than imposition). The agents of transfer are cognitively dominant in the RL Khuzestan Arabic, the agents’ L1. Even though Persian is generally acquired during childhood and today is spoken by most speakers, it usually is the speakers’ L2. Cases of convergence (cf. Lucas 2015: 530–531) are possible in the present contact situation among speakers with a very high (L1-like) command of Persian, for example university students. But of course it is hard to draw an exact line between L1 and L2 proficiency and thus between convergence and borrowing (cf. Lucas 2015: 531).

3.2 Phonology

As in other Bedouin Arabic dialects, the presence of the phonemes /č/ and /g/ is ultimately the result of internal development from original *k and *q, rather than borrowing from Persian (see Procházka, this volume).

The phoneme /p/, e.g. *perde* ‘curtain’ < Pers *parde*, is also common in all Iraqi dialects and probably emerged in this region due to contact with Persian and Kurdish (see Procházka, this volume).

An interesting phonological feature of Khuzestan Arabic is that /ɣ/ often reflects etymological *q,⁷ which is otherwise realized as /g/ and /ğ/. It is most likely

⁵The data used for the present analysis was collected mainly in Aḥwāz, Muḥammara (Khorramshahr), Ḥamīdiyye and Ḥafaḡiyye (Susangerd) in 2016. The male and female informants were bilingual as well as monolingual Khuzestan Arabic speakers from 25 to over 70 years old.

⁶Sakel (2007: 15) defines matter replication as the replication of “morphological material and its phonological shape”.

⁷This phenomenon is also documented for the Arabic dialects of Kuwait, Qatar, and the United Arab Emirates (Holes 2016: 54, fn. 5).

1 *Khuzestan Arabic*

that the shift /y/ < *q first occurred in Persian borrowings ultimately of Arabic origin, e.g. *yisma* ‘part, section’, *taşdıy* ‘driving licence’, *yifruy* ‘it differs’, *tayriban* ‘approximately’, and *yidra* ‘power’. This feature is most probably a transfer from Persian,⁸ in which both *q and *y in Arabic loanwords are always pronounced /y/ (Matras & Shabibi 2007: 138).⁹ Later, this phonological change further affected lexemes which have no cognate form in Persian, e.g. *bayra* ‘cow’, a borrowing from Modern Standard Arabic (the KhA dialectal form being *hāyša* ‘cow’). There are, however, certain lexemes which are not affected by this rule, e.g. *gāl* ‘he said’. Other lexemes show free variation in the pronunciation of /g/, e.g. *gabul/yabul* ‘formerly, before’.

Lexical borrowings are often adapted to Arabic phonology. For example, Khuzestan Arabic speakers of the older generation usually pronounce the phoneme /p/ as [b]; and /s/ in Persian loanwords is sometimes pronounced [š], e.g. *berde* ‘curtain’ < Pers *parde*, and *šebzi* ‘herb stew’ < Pers *sabzi*.

Negative structures bear stress on the first syllable,¹⁰ e.g. KhA *mā arūḥ* ‘I don’t go’. This is a feature shared with some Persian and Turkish varieties and other North East Arabian dialects (Ingham 2005: 178–179). This common phonological characteristic therefore seems to be a Sprachbund phenomenon of the Mesopotamian region, which reflects the long history of contact and migration across language boundaries due to trade, war, shared cultural practices, nomadism, etc. (Winford 2003: 70–74). Though the directions and mechanisms of borrowing within the languages of a Sprachbund are often hard to categorize (Winford 2003: 74), we can probably assume that Khuzestan Arabic, being spoken by a minority group, has borrowed and adapted this phonological stress pattern under RL agentivity.

3.3 Morphosyntax

3.3.1 Replication of Persian phrasal verbs

The replication of phrasal verbs is a contact phenomenon also found in the Arabic varieties of Turkey (Grigore 2007: 157–159; and Procházka, this volume). As shown in examples (1)–(4), Khuzestan Arabic replicates Persian phrasal verbs by substituting the Persian light verbs with Khuzestan Arabic equivalents and directly

⁸Contrast Holes (2016: 53–54), who explains the /y/-/q/ merger among the Najd-descendent Bahraini Arabic speakers as internal development.

⁹In Modern Persian the phoneme /y/ has two allophones, [c] and [ɣ] (Majidi 1986: 58–60).

¹⁰Ingham (1991: 724) describes this phenomenon also for Khuzestan Arabic *wh*-interrogatives and prepositions.

Bettina Leitner

replicating the Persian nouns (cf. [Matras & Shabibi 2007](#): 142). The noun in example (1) is Arabic in its origins but its usage in a phrasal verb construction with a new meaning is a Persian innovation.

- (1) a. Aḥwāz, Khuzestan, male, 26 years (own data)
 ṭəgg muḥḥ
 hit.PRF.3SG.M brain
- b. Persian
 muḥḥ zadan
 brain hit.INF
 ‘to brainwash/convince someone’¹¹
- (2) a. Aḥwāz, Khuzestan, male, 39 years (own data)
 kaḏḏ īrād
 take.PRF.3SG.M nagging
- b. Persian
 īrād gereftan
 nagging take.INF
 ‘to pick at someone’

As examples (3) and (4) show, the Persian nouns are sometimes adapted morphophonologically.

- (3) a. Aḥwāz, Khuzestan, male, 50 years (own data)
 sawwa ʔōmāde
 make.PRF.3SG.M ready
- b. Persian
 āmāde kardan
 ready make.INF
 ‘to prepare sth.’
- (4) a. Aḥwāz, Khuzestan, male, 26 years (own data)
 ṭalaʔ yabūli¹²
 emerge.PRF.3SG.M acceptance
- b. Persian
 yabūl šodan
 acceptance become.INF
 ‘to pass (an exam), be accepted’

¹¹ All Persian translations are given in the modern spoken (Tehrani) variety of Persian.

The pattern for phrasal verbs – transferred into the RL Khuzestan Arabic under RL agentivity – provides Khuzestan Arabic with an easy way to convert foreign nouns into verbs.

As illustrated in examples (5) and (6), the pattern is adapted according to Arabic syntactic rules: (i) the verb is moved into the initial position; and (ii) a direct object is introduced between verb and nominal element (post-verbally). In Persian, however, the verb always remains in final position following the nominal element and a direct object would be introduced before both elements (see e.g. Majidi 1990: 447–448).

- (5) a. Aḥwāz, Khuzestan, male, 26 years (own data)
 ṭaggi yāndart-i wāks
 hit.IMP.2SG.F shoe-OBL.1SG wax
- b. Persian
 kaḡš-am-o wāks be-zan
 shoe-OBL.1SG-ACC wax IMP-hit.PRS
 ‘Polish my shoes!’
- (6) a. Aḥwāz, Khuzestan, female, 35 years (own data)
 yṭaggūn aṭ-ṭamāṭe rande
 hit.IMP.3PL.M DEF-tomato grater
- b. Persian
 gūḡe_farang-i rō rande mī-zanan
 tomato-INDF ACC grater PROG-hit.PRS.3PL.M
 ‘They grate some tomato.’

This structure has become productive in Khuzestan Arabic. For example, in the phrasal verb *ṭagg dabbe* ‘to cheat’ (lit. ‘to hit a water canister’) both the verb and noun are taken from Khuzestan Arabic and only the construction’s syntactic pattern is taken from Persian.

3.4 Syntax

3.4.1 Definiteness marking

Matras & Shabibi (2007: 141–142) see Khuzestan Arabic relative clauses without definite heads as evidence for the decline of overt definiteness marking in Khuze-

¹²The final *-i* in *yabūli* probably originates from the Persian indefiniteness marker *-i* (see Majidi 1990: 309–314) and has become part of this word in Khuzestan Arabic, so that *yabūli* is monomorphemic.

Bettina Leitner

stan Arabic, based on the Persian model with generally unmarked definiteness, e.g. *mara lli šiftū-ha ḥābarat* ‘The woman that you saw called’ (2007: 142). However, this pattern is also documented in Arabic dialects which have had no contact with Persian (Pat-El 2017: 454–455; cf. Procházka 2018: 269).

Matras & Shabibi (2007: 140) further postulate that the Persian *ezāfe* pattern in adjectival attribution is replicated in Khuzestan Arabic.¹³ According to their theory, the construct state marker *-t* (with an indefinite head) and/or the definite article (of the attribute) are reanalysed as markers of attribution matching the Persian *ezāfe* marker *-(y)e*, as in (7).

- (7) a. KhA (Matras & Shabibi 2007: 140)
 ḡazīra-t l-ḥaḍra
 island-CON DEF-green.F
- b. Persian
 ḡazīre-ye sabz
 island-EZ green
 ‘the green island’

However, this pattern is also observed in other modern Arabic dialects which have not been exposed to Persian influence as well as in older forms of Arabic.¹⁴

Consequently, it is highly unlikely that this phenomenon has developed due to Persian influence, although it cannot be ruled out that contact with Persian has fostered the preservation of this apparently old feature.

3.4.2 Word order changes

Khuzestan Arabic shows no changes due to contact in basic word order.¹⁵ The only attested word order changes concern the position of the verbs *čān* ‘to be’ and *šār* ‘to become’, both of which can appear in final position as an unmarked

¹³See e.g. Ahadi (2001: 103–109) for the usage of the Persian *ezāfe*.

¹⁴See Pat-El (2017: 445–449) for numerous examples from different modern Arabic dialects, Middle Arabic, Pre-Islamic Arabic, Quran Arabic, and other Central Semitic languages. See also Retsö (2009: especially 21–22) and Procházka (2018: 267–269), who also proves that this is an old feature already found in Old Arabic and points out that it is mainly found among dialects which are spoken in regions with no or only marginal influence from Modern Standard Arabic.

¹⁵Ingham (1991: 715) states that in Khuzestan Arabic neither VSO nor SVO word order is particularly dominant. Matras & Shabibi (2007: 147) postulate that the usage of OV order in Khuzestan Arabic is increasing as “the beginning of a shift in word order” on the basis of the Persian type, where OV prevails. In both of their examples the objects are topicalized (with pronominal resumption), which is a common phenomenon in spoken Arabic (Brustad 2000: 330–333; 349), and as such not obviously the result of Persian influence.

construction. This sentence-final position in no case functions as the default, and is in fact less frequent than its non-final position.¹⁶ *čān* or *šār* in final position are never stressed.

The sentence-final position of *čān* or *šār* (see examples (8)–(10)) is likely a pattern replication of the Persian model, i.e. sentences with final *būdan* ‘to be’ or *šodan* ‘to become’.

- (8) a. ʕAbbādān, Khuzestan, male, 35 years (own data)
 šuyul-hum b-əl-bandar čān
 work-3PL.M in-DEF-port be.PRF.3SG.M
 b. Persian
 kār-ešūn tū-ye bandar būd
 job-OBL.3PL in-EZ port be.PST.3SG
 ‘Their job was at the port.’
- (9) a. Muḥammara, Khuzestan, male, 30 years (own data)
 əğdād-i mallāk-in čānaw
 grandparents-OBL.1SG owner-PL be.PRF.3PL.M
 b. Persian
 pedarbozorgā-m mālek būdan
 grandparents-OBL.1SG owner be.PST.3PL
 ‘My grandparents were owners [of land].’
- (10) a. Aḥwāz, Khuzestan, female, 40 years (own data)
 hassa šway l-māy bārəd šār
 now a_bit DEF-water cold become.PRF.3SG.M
 b. Persian
 alʔān yekam ʔāb sard šod
 now a_bit water cold become.PST.3SG
 ‘The water has become a bit cold now.’

The next example might show a tendency to use a present-tense copula with human subjects, expressed with the verb *šār* ‘to become’:

- (11) a. Aḥwāz, Khuzestan, female, 35 years (own data)
 əhya mart uḥū-y əşşīr
 3SG.F wife brother-OBL.1SG COP.IMP.F.3SG.F

¹⁶In my data, *čān* appears 23 of 152 times in sentence-final position, *šār* 11 of 165 times. The additional examples are taken from my questionnaire.

Bettina Leitner

b. Persian

ūn zan dādāš-am-e

3SG wife brother-OBL.1SG-COP.PRS.3SG

‘She is the wife of my brother.’

In the Khuzestan Arabic construction for pluperfect tense, *čān* can also appear in sentence-final position, after the active participle. This construction, although not very frequent, is very likely a direct transfer of the Persian structure, in which the auxiliary *būdan* also follows the participle.¹⁷

(12) a. Aḥwāz, Khuzestan, male, 26 years (own data)

lamman əyəna l-əl-bīət, əhma mākl-īn čānaw

when come.PRF.1PL to-DEF-house 3PL.M eat.PTCP-PL.M be.PRF.3PL.M

b. Persian

vayti-ke mā bargāštīm ḥūne, ūnhā yazā-ro ḥorde

when-REL 1PL come_back.PST.1PL home 3PL food-ACC eat.PTCP

būdan

be.PST.3PL

‘When we came home, they had (already) eaten.’

This word order change has probably been triggered by the high frequency in speech of Persian sentences with forms of *būdan* in final position. Lucas (2012: 295) explains the usage of foreign patterns as the result of the human cognitive tendency to minimize the high processing efforts associated with the extensive use of two languages.¹⁸

čān is also used in sentence-final positions after the main verb in the imperfect in Khuzestan Arabic constructions expressing the continuous past. In spoken Persian, the continuous past is formed without a sentence-final *būdan*.¹⁹ This

¹⁷Matras and Shabibi (2007: 142–143) describe the use of this construction as a change in the Khuzestan Arabic tense system. However, the pattern *kān* + active participle is also commonly used in other Arabic dialects to express pluperfect meaning or to describe completed actions which have an impact on the present, see for example Denz (1971: 92–94; 115–116) for Iraqi (Kwayriš) and Grotzfeld (1965: 88) for Syrian Arabic.

¹⁸Connections between units of a neural network associated with certain syntactic patterns can be strengthened from repeated exposure to and use of that pattern (Lucas 2012: 291). Hence, the employment of a Persian syntactic structure in Khuzestan Arabic needs less processing effort because the same strengthened neural network is activated.

¹⁹The Persian continuous past is formed with the particle *mī* prefixed to the simple past of the respective main verb and can (for the progressive form) be preceded by the simple past of *dāštan* ‘to have’: e.g. (*dāšt*) *mī-raft* ‘he was going’ (Majidi 1990: 232, 235).

1 Khuzestan Arabic

case is not a direct transfer of the Persian pattern, but perhaps a construction analogous to the pluperfect and other Persian forms with *būdan* in final position.

- (13) a. Aḥwāz, Khuzestan, male, 55 years (own data)
 hāda ham mən zuḡur yəštəḡəl čān
 DEM.SG.M also from childhood work.IMPF.3SG.M be.PRF.3SG.M
 b. Persian
 īn ham az kūdeki kār mī-kard
 DEM.SG also from childhood work PROG-do.PST.3SG
 ‘This one has also been working from childhood on.’

Example (14) shows both syntactic variants in one sentence, i.e. *čān* before and after the main verb.

- (14) a. Muḥammara, Khuzestan, female, 40 years (own data)
 umm-i čānat təṭṭaḡḡab, eh, əb-zamān əš-šāh,
 mother-OBL.1SG be.PRF.3SG.F veil.IMPF.3SG.F yes in-time DEF-shah
 bass tətḡawwəš čānat
 only veil.IMPF.3SG.F be.PRF.3SG.F
 b. Persian
 mādar-am (dāšt) naqāb mī-zad, āre, dar zamān-e
 mother-1SG (have.PST.3SG) veil PROG-hit.PST.3SG yes in time-EZ
 šāh, hamīše naqāb mī-zad
 shah always veil PROG-hit.PST.3SG
 ‘My mother was veiling her face (with a *būšiyye*),²⁰ yes, during the times of the shah, she was always veiling her face.’

Because all the above examples equally work with *čān/šār* in non-final position, this process of word-order-related pattern replication in Khuzestan Arabic is still ongoing. Indeed, all informants, when asked for the correct structure in the above examples, preferred the verb *čān* in non-final position.²¹

Lucas (2015: 530–531) explains the basic word order changes (from VSO to SOV) in Bukhara Arabic (cf. Ratcliffe 2005: 143–144; and Versteegh 2010: 639) as

²⁰ *būšiyye* or *pūšiyye* ‘veil’ is also documented for Iraqi Arabic (Woodhead & Beene 1967: 53).

²¹ My informants from Baghdad considered all constructions with *čān* in final position to be wrong. However, this structure is used in Baṣra Arabic (Qasim Hassan, personal communication, January 2018).

Bettina Leitner

a result of convergence with Uzbek.²² Although a clear division between convergence and borrowing is hard to make, I consider the contact-induced word order changes that occur in Khuzestan Arabic to be instances of borrowing because most speakers are clearly native speakers of, and therefore dominant in, Khuzestan Arabic only.

3.4.3 *ḥōš* preceding verbs and nouns

In Persian, *ḥōš* ‘good, well’ is used as a prefixed (lexicalized) element preceding some nouns and verbs to coin compound adjectives, nouns, and verbs (Majidi 1990: 411, 413): e.g. Pers *ḥōš-andām* ‘handsome’ (< *andām* ‘shape; body’), *ḥōš-nevīs* ‘calligrapher’ (< present stem *nevīs-* ‘to write’), and *ḥōš-āmadan* ‘to please’ (< *āmadan* ‘to come’).

Khuzestan Arabic has borrowed some of these Persian compound adjectives: e.g. *ḥōš-bū* ‘nice-smelling’ (< Pers *bū* ‘smell, scent’), *ḥōš-tīp* ‘handsome’ (< Pers *tīp* ‘type’), and *ḥōš-aḥlāq* ‘(with) good manners’ (< Pers *aḥlāq* ‘decency; ethics, morality’, pl. of *ḥolq* ‘character, nature’). However, in Khuzestan Arabic the use of the Persian element *ḥōš* has been further developed. It is productively used as an attributive adjective preceding nouns, but not agreeing in gender or number with them, e.g. *ḥōš walad* ‘a good boy’, *ḥōš abnayya* ‘a good girl’, *ḥōš banāt* ‘good girls’, *ḥōš əwlād* ‘good kids’.²³

3.5 Lexicon

3.5.1 Lexical transfer

The greatest influence from Persian on Khuzestan Arabic has occurred in lexicon. Many Persian lexemes were borrowed generations ago. The most frequently borrowed elements are nouns denoting cultural or technological innovations which have filled lexical gaps in the RL Khuzestan Arabic. Verbs, adverbs, adjectives, and many discourse particles have also been borrowed from the SL Persian.

The majority of the examples below are cases of transfer of morpho-phonological material (matter) and semantic meaning (pattern) under RL agentivity.

Many of the Persian borrowings have been phonologically and morphologically integrated into the RL. For instance, for many borrowed Persian nouns

²²Lucas (2015: 525) defines convergence as changes made to a language under the agentivity of speakers who are native speakers of both the SL and the RL.

²³This construction is also found in Iraqi Arabic (cf. Erwin 1963: 256), which might prove that the element *ḥōš* is an older borrowing.

1 Khuzestan Arabic

Arabic internal plural forms are created, e.g. *ḥatākīr* ‘ball-point pens’ (sg. *ḥatkār* < Pers *ḥod-kār* ‘ball-point pen’), or *banādār* ‘ports’ (sg. *bandar* < Pers *bandar* ‘port’).

Again, the borrowing of foreign (L2) elements into the speakers’ L1 might be explained by the human cognitive tendency to minimize the processing effort in lexical selection between two languages (Lucas 2012: 291; see §3.4.2). So if a certain Persian word is frequently used and often heard (for example at school), the connections of a neural network associated with this word are strengthened (Lucas 2012: 291), which makes it easier to employ the word in one’s L1.

3.5.2 Semantic fields

The following illustrative list of Persian loans in Khuzestan Arabic shows the most important semantic fields of lexical borrowing.

- Administration/military: *farmāndāri* ‘governorship’ < Pers *farmāndāri*; *čārra* ‘crossroad’ < Pers *čahār-rāh*; *sarbāz*/šarbāz ‘soldier’ < Pers *sarbāz*.
- Agriculture: *šalafkoš* ‘pesticide’ (lit. weed-killer) < Pers *šalaf-koš*; *kūd* ‘dung’ < Pers *kūd*.
- Dress/textiles: *dāmen* ‘skirt’ < Pers *dāman*; *šāla* ‘head covering’ < Pers *šāl* ‘Kashmir shawl’ (Ingham 2005: 174).
- Education: *klāš* ‘class, grade’ < Pers *kelās*; *ḥatkār* ‘biro’ < Pers *ḥod-kār*; *dānišga* ‘university’ < Pers *dānišgāh*.
- Food: *ğālfari* ‘parsley’ < Pers *ğālfari*; *češmeš* ‘raisins’ < Pers *kešmeš*; *šalyam* ‘turnip’ < Pers *šalyam*; *serke* ‘vinegar’ < Pers *serke*.
- Material culture: *šīše* ‘bottle’ < Pers *šīše*; *ğām* ‘(window) glass’ < Pers *ğām*; *tīye* ‘blade’ < Pers *tīye*; *yəḥčāle* ‘refrigerator’ < Pers *yahčāl*; *sīm buksel* ‘towrope’ < Pers *sīm-e boksol*; *perde/berde* ‘curtain’ < *parde*; *gīre* ‘hair barrette’ < Pers *gīre-ye sar/mūy*; *mīz* ‘table’ < Pers *mīz*; *dariše* ‘window’ < Pers *dariče*; *pəṅğara* ‘window’ < Pers *panğare*.
- Other: *yīme* ‘price’ < Pers *yīmat*; *bandar* ‘port’ < Pers *bandar*; *nāmard* ‘brute’ < Pers *nāmard*.

Some items ultimately of Arabic origin have been re-borrowed into Khuzestan Arabic from Persian, preserving the Persian meaning, e.g. KhA *bəryi* ‘electronic’ < Pers *bary* ‘electricity; lightning’ < Arabic *barq* ‘lightning’.

Bettina Leitner

3.5.3 Verbs and adverbs

Khuzestan Arabic verbs and adverbs resulting from language contact are always morphologically integrated. These are either directly borrowed Persian verbs, e.g. *bannad* ‘to close (e.g. the tap)’ < Pers imperfect and present stem *band-* ‘close’;²⁴ *gayyār* ‘to get stuck’ < Pers *gīr kardan* ‘to get stuck’; *šammār* ‘to repair’ < Pers *tašmīr kardan* ‘to repair’; *čassāb* ‘to glue’ < Pers *časb zadan* ‘to glue’; *gəzar* ‘to pass (time)’ < Pers present stem *gozar-* ‘to pass (time)’ (see example (15) below);²⁵ or Persian nouns turned into Khuzestan Arabic (ad)verbs, e.g. *əb-zūr* ‘by force’ < Pers *zūr* ‘power; violence; force’; and *zaḥəm* ‘to bother’ (transitive) < Pers *zahmat* ‘trouble, bother’ (see examples (16) and (17) below).²⁶

- (15) Aḥwāz, Khuzestan, male, 26 years (own data)
 čā hāy əl-ḥayāt lō la? təgzar baʔad, təmši
 DM DEM.F DEF-life or no pass.IMPF.3SG.F after_all go.IMPF.3SG.F
 ‘See, that is how life is, right? It passes by (quickly), it goes.’
- (16) Aḥwāz, Khuzestan, male, 26 years (own data)
 zaḥmīət-kum, ʔafwan
 bother.PRF.1SG-2PL.M sorry
 ‘Sorry, I must have bothered you/caused you trouble.’²⁷
- (17) KhA, Aḥwāz (own data)
 mumkin azaḥm-ək əb-šuyla
 possible bother.IMPF.1SG-2SG.M with-issue
 ‘May I bother you with something/ask you a favour?’

3.5.4 Discourse elements

A range of Persian discourse elements have been borrowed by Khuzestan Arabic (cf. [Matras & Shabibi 2007](#): 143–145),²⁸ e.g. KhA *ham/hamme* ‘also, as well’ <

²⁴ Also common in the Gulf region and in Yemen ([Behnstedt & Woidich 2014](#): 290).

²⁵ The verb *gəzar* is used only in phrases that refer to the “passing by” of life.

²⁶ Khuzestan Arabic *zaḥme* is also used as a noun for a rebuke, e.g. *zaḥme šalīāk!* ‘Shame on you!’, which would be expressed in a different way in Persian: *heḡālat nemi keši?* ‘Shame on you!’ (lit. ‘Are you not ashamed?’).

²⁷ A phrase often used when leaving, for example after an invitation for dinner.

²⁸ [Matras & Shabibi \(2007: 144\)](#) claim that the Persian conjunctions *agarče* and *bāinke*, both meaning ‘although/even though’, and the Persian factual complementizer *ke* ‘that’ have also been borrowed by Khuzestan Arabic. However, I have found no evidence for their usage in my data.

1 Khuzestan Arabic

Pers *ham* and KhA *ham...ham* ‘(both)...and’ < Pers *ham...ham*;²⁹ and KhA *hič* ‘nothing; no(t)... at all’ < Pers *hič*.³⁰

The Khuzestan Arabic discourse elements *hō*/*hōš* ‘well; okay’ < Pers *hō(b)*/*hōš* are often used phrase-initially, (18).³¹ They are of Persian origin, but have partly adopted a different form and function in Khuzestan Arabic.³²

- (18) Aḥwāz, Khuzestan, male, 55 years (own data)
hōš, š-ʔəd-na, taʔay əhna baba
 DM what-at-1PL come.IMP.SG.F here father
 ‘Okay, what (else) do we have, come here, dear!’

Both *hō* and *hōš* are also often used in stories following the verb *gāl* ‘to say’.

- (19) Aḥwāz/Fəllāḥiyya, Khuzestan, female, 50 years (own data)
lamman yada mən ʔəd-hum, gāl-la hō,
 when leave.PRF.3SG.M from at-3PL.M say.PRF.3SG.M-DAT.3SG.M DM
hāy ər-rummānāt š-asawwi bī-hən
 DEM.SG.F DEF-pomegranate.PL what-make.IMP.F.1SG with-3PL.F
 ‘When he left them, he said to him, “Well, what shall I do with these pomegranates?”’

4 Conclusion

Because of the dominance of Persian in the Iranian educational system and work environment, the lack of influence from Modern Standard Arabic, and the long period of geographical proximity, the Persian-speaking society of southwest Iran has left many linguistic traces in the language of the Arabic-speaking community of Khuzestan.

²⁹This discourse element is also known for Iraq (Malaika 1963: 36) and, like KhA *hast/hassət* ‘there is’ < Pers *hast* (Ingham 1973: 25, fn.27), is probably an older borrowing.

³⁰Shabibi (2006: 176–177) derives KhA *balkət* ‘maybe, hopefully’ from Pers *balke ham*, which can mean ‘maybe’. A Turkish origin of this word seems more likely: cf. Aksoy (1963: 620) for the existence of *belke/belkit* in Eastern Turkish dialects. Malaika (1963: 35) also derives the Baghdadi Arabic *belki* ‘rather, maybe’ from Turkish, as does Seeger (2009: 28) for *balki, balkiṣ, balkin* ‘maybe; possibly; probably’ in Ramallah Arabic.

³¹According to my informants and data, the form *hōb* is not used in Khuzestan Arabic (contrast Matras & Shabibi 2007: 143).

³²In Persian, *hō(b)* is a discourse particle and simple adjective, and *hōš* just an adjective (see §3.4.3; Shabibi 2006: 160). Thus the Persian adjective *hōš* has been desemanticized in Khuzestan Arabic to function as a discourse particle with the meaning ‘well, okay’ (Shabibi 2006: 160).

Bettina Leitner

Van Coetsem (2000: 59; cf. Lucas 2015: 532) suggests that lexical, but not syntactic and phonological transfer is to be expected under RL agentivity. However, Khuzestan Arabic phonology and syntax have been influenced by the SL Persian under RL agentivity, albeit to a much lesser extent than the lexicon.

Khuzestan Arabic does not show transfer of patterns from Persian in either inflectional or derivational morphology. However, we do find an adapted pattern replication of Persian phrasal verbs (with preservation of the Arabic word order).

As for syntax and contact-induced word order changes, the alternative sentence construction with *čān* in sentence-final position can be explained as a result of Persian influence on Khuzestan Arabic. This change might have been triggered by the similar and very frequent Persian constructions with sentence-final *būdan*. Thus, we do have some syntactic change due to transfer under RL agentivity, which van Coetsem considered to be unexpected (see above).

Persian lexical items have often been borrowed in Khuzestan Arabic for novel concepts (lexical gaps), which is why semantic fields relating to technical or cultural innovations, education, and administration show the greatest amount of Persian borrowing. This also explains why nouns are generally more often transferred than verbs (cf. Lucas 2015: 532). Persian words are regularly integrated into Khuzestan Arabic phonology and morphology, for example the Arabic internal plural is formed for Persian nouns. Also, many discourse particles have been transferred from Persian into Khuzestan Arabic. Some of them, e.g. *ham* ‘also’, had been in use generations ago among Arabic speakers in Khuzestan and beyond (Iraq, Gulf).

Of course, contact between Khuzestan Arabic and Persian has always been limited to certain social contexts (outside the family), especially for women, who had and still have much less access to education and employment and thus to the Persian-speaking world. This fact, and some structural differences between the languages, explain the limits of contact-induced language change in Khuzestan Arabic, especially in morphology and syntax.

Hopefully, future research on the dialects of Khuzestan will provide more empirical data on instances of contact-induced change. An enlarged database should especially provide further evidence concerning the development and extent of word order changes.

Further reading

Ingham (2011) provides a sketch grammar of Khuzestan Arabic.

Ingham (2005) discusses Turkish and Persian borrowings in Khuzestan Arabic

and northeastern Arabian dialects.

Matras & Shabibi (2007) is an article on contact-induced changes in Khuzestan Arabic based on Shabibi (2006).

Shabibi (2006) is an unpublished doctoral dissertation on contact-induced changes in Khuzestan Arabic.

Acknowledgements

I would like to thank Stephan Procházka and Dina El-Zarka for their critical remarks and bibliographical suggestions on this chapter. I would additionally like to thank my informant and good friend Majed Naseri for all his help on the transcription and translation of my recordings.

Abbreviations

ACC accusative	KhA Khuzestan Arabic
COP copula	M masculine
DAT dative	OBL oblique
DEF definite	Pers Persian
DEM demonstrative	PL/pl. plural
DM discourse marker	PTCP participle
EZ Persian <i>ezāfe</i>	PRF perfect
F feminine	PROG progressive
IMP imperative	PRS present
IMPF imperfect	PST past
INF infinitive	REL relative particle
INDF indefinite	SG/sg. singular

References

- Ahadi, Shahram. 2001. *Verbergänzungen und zusammengesetzte Verben im Persischen: Eine valenztheoretische Analyse*. Wiesbaden: Reichert.
- Aksoy, Ömer A. (ed.). 1963. *Türkiyede halk ağzından derleme sözlüğü*. Ankara: Türk Tarih Kurumu Basımevi.
- Behnstedt, Peter & Manfred Woidich. 2014. *Wortatlas der arabischen Dialekte III: Verben, Adjektive, Zeit und Zahlen*. Leiden: Brill.

Bettina Leitner

- Brustad, Kristen. 2000. *The syntax of spoken Arabic: a comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Georgetown: Georgetown University Press.
- Daniel, E. L. 1986. Arab settlements in Iran. In Ehsan Yarshater (ed.), *Encyclopaedia Iranica*, vol. 2, 210–214. London: Routledge & Kegan Paul.
- De Planhol, X. 1986. Ābādān, ii. Modern Ābādān. In Ehsan Yarshater (ed.), *Encyclopaedia Iranica*, vol. 2, 53–57. London: Routledge & Kegan Paul.
- Denz, Adolf. 1971. *Die Verbalsyntax des neuarabischen Dialektes von Kwayriš (Irak): Mit einer einleitenden allgemeinen Tempus- und Aspektlehre*. Wiesbaden: Steiner.
- Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic*. Washington: Georgetown University Press.
- Gazsi, Dénes. 2011. Arabic–Persian language contact. In Stefan Weninger, Janet C. E. Watson, Michael Streck & Geoffrey Khan (eds.), *The Semitic languages: An international handbook*, 1015–2021. Berlin: De Gruyter Mouton.
- Grigore, George. 2007. *L’arabe parlé à Mardin: Monographie d’un parler arabe “périphérique”*. Bucharest: Editura Universităţii din Bucureşti.
- Grotzfeld, Heinz. 1965. *Syrisch-arabische Grammatik (Dialekt von Damaskus)*. Wiesbaden: Harrassowitz.
- Holes, Clive. 2016. *Dialect, culture, and society in eastern Arabia*. Vol. 3: *Phonology, morphology, syntax, style*. Leiden: Brill.
- Ingham, Bruce. 1973. Urban and rural arabic in Khūzistān. *Bulletin of the School of Oriental and African Studies* 36. 533–553.
- Ingham, Bruce. 1976. Regional and social factors in the dialect geography of southern Iraq and khuzistan. *Bulletin of the School of Oriental and African Studies* 39. 62–82.
- Ingham, Bruce. 1991. Sentence structure in khuzistani Arabic. In Alan S. Kaye (ed.), *Semitic studies: In honor of Wolf Leslau on the occasion of his eighty-fifth birthday, November 14th, 1991*, vol. 1, 714–728. Wiesbaden: Harrassowitz.
- Ingham, Bruce. 2005. Persian and Turkish loans in the Arabic dialects of North Eastern arabia. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: case studies from iranian, semitic and turkic*, 173–180. London: Routledge.
- Ingham, Bruce. 2011. Khuzestan Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Lucas, Christopher. 2012. Contact-induced grammatical change: Towards an explicit account. *Diachronica* 29. 275–300.

1 Khuzestan Arabic

- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Majidi, Mohammad-Reza. 1986. *Strukturelle Grammatik des Neupersischen (Fārsi). Band I: Phonologie*. Hamburg: Buske.
- Majidi, Mohammad-Reza. 1990. *Strukturelle Grammatik des Neupersischen (Fārsi). Band II: Morphologie*. Hamburg: Buske.
- Malaika, Nisar. 1963. *Grundzüge der Grammatik des arabischen Dialektes von Bagdad*. Wiesbaden: Harrassowitz.
- Matras, Yaron & Maryam Shabibi. 2007. Grammatical borrowing in khuzistani Arabic. In Yaron Matras, Georg Bossong & Bernard Comrie (eds.), *Grammatical borrowing in cross-linguistic perspective*, 137–149. Berlin: Mouton de Gruyter.
- Oberling, P. 1986. Arab tribes of Iran. In Ehsan Yarshater (ed.), *Encyclopaedia Iranica*, vol. 2, 215–219. London: Routledge & Kegan Paul.
- Pat-El, Na'ama. 2017. Digging up archaic features: Neo-Arabic and comparative semitics in the quest for proto-Arabic. In Ahmad Al-Jallad (ed.), *Arabic in context: Celebrating 400 years of Arabic at Leiden University*, 441–475. Leiden: Brill.
- Procházka, Stephan. 2018. The northern fertile crescent. In Clive Holes (ed.), *Arabic historical dialectology: linguistic and sociolinguistic approaches*, 257–292. Oxford: Oxford University Press.
- Ratcliffe, Robert. 2005. Bukhara arabic: a metatypized dialect of arabic in central asia. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: case studies from iranian, semitic and turkic*, 141–51. London: Routledge.
- Retsö, Jan. 2009. Nominal attribution in Semitic: Typology and diachrony. In Janet C. E. Watson & Jan Retsö (eds.), *Relative clauses and genitive constructions in Semitic*, 3–33. Oxford: Oxford University Press.
- Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In Yaron Matras, Georg Bossong & Bernard Comrie (eds.), *Grammatical borrowing in cross-linguistic perspective*, 15–29. Berlin: Mouton de Gruyter.
- Seeger, Ulrich. 2009. *Der arabische Dialekt der Dörfer um Ramallah. Teil 2: Glossar*. Wiesbaden: Harrassowitz.
- Shabibi, Maryam. 2006. *Contact-induced grammatical changes in Khuzestani Arabic*. Manchester: University of Manchester doctoral dissertation.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.

Bettina Leitner

- Versteegh, Kees. 2010. Contact and the development of Arabic. In Raymond Hickey (ed.), *The handbook of language contact*, 634–651. Oxford: Wiley-Blackwell.
- Winford, Donald. 2003. *An introduction to contact linguistics*. Malden: Blackwell.
- Woodhead, D. R. & Wayne Beene. 1967. *A dictionary of Iraqi Arabic (Arabic–English)*. Washington D.C.: Georgetown University Press.

Chapter 2

Anatolian Arabic

Faruk Akkuş

University of Pennsylvania

This chapter investigates contact-induced changes in Anatolian Arabic varieties. The study first gives an overview of the current state and historical development of Anatolian Arabic. This is followed by a survey of changes Anatolian Arabic varieties have undergone as a result of language contact with primarily Turkish and Kurdish. The chapter demonstrates that the extent of the change varies from one dialect to another, and that this closely correlates with the degree of contact a dialect has had with the surrounding languages.

1 Current state and historical development

Anatolian Arabic is part of the so-called *qəltu*-dialect branch of the larger Mesopotamian Arabic, and essentially refers to the Arabic dialects spoken in eastern Turkey.¹ In three provinces of Turkey – Hatay, Mersin and Adana – Syrian sedentary Arabic, is spoken (see Procházka, this volume, for discussion of these dialects). Other than these dialects, in Jastrow's (1978) classification of Mesopotamian *qəltu* dialects, Anatolian Arabic dialects are subdivided into five groups: Diyarbakır dialects (spoken by a Jewish and Christian minority, now almost extinct); Mardin dialects; Siirt dialects; Kozluk dialects; and Sason dialects. In his later work, Jastrow (2011a) classifies Kozluk and Sason dialects under one group along with Muş dialects – investigated primarily by Talay (2001; 2002). The two larger cities where Arabic is spoken are Mardin and Siirt, although in the latter Arabic is gradually being replaced by Turkish.

¹This group represents an older linguistic stratum of Mesopotamia as compared to the *gəlat* dialects. The terms *qəltu* vs. *gəlat* dialects are due to Blanc (1964), who distinguished between the Arabic dialects spoken by three religious communities, Muslim, Jewish, and Christian, in Baghdad. He classified the Jewish and Christian dialects as *qəltu* dialects and the Muslim dialect as a *gəlat* dialect, on the basis of their respective reflexes of Classical Arabic *qultu* 'I said'.



Faruk Akkuş

The linguistic differences between these various Arabic-speaking groups are quite considerable. Thus, given the low degree of mutual intelligibility, speakers of different varieties resort to the official language, Turkish, to communicate. Jastrow (2006) reports an anecdote, wherein high school students from Mardin and Siirt converse in Turkish, since they find it difficult to understand each other's dialects. Expectedly, mutual intelligibility is at a considerably higher level among different varieties of a single group, despite certain differences. For instance, speakers of Kozluk and Muş Arabic have no difficulty in communicating with one another in Arabic.

The existence of Anatolian dialects closely relates to the question of the arabization of the greater Mesopotamian area. Although the details largely remain obscure, a commonly-held view is that it took place in two stages: the first stage concerns the emergence of urban varieties of Arabic around the military centers, such as Baṣra or Kūfa, during the early Arab conquests. Later, the migration of Bedouin dialects of tribes added another layer to the urban dialects (see e.g. Blanc 1964; Versteegh 1997; Jastrow 2006 for discussion). According to Blanc (1964), the *qeltu* dialects are a continuation of the medieval vernaculars that were spoken in the sedentary centers of ʿAbbāsīd Iraq. Blanc (1964) also noted that the *qeltu* dialects did not stop at the Iraqi–Turkish border, but in fact continued into Turkish territory. He mentioned the towns of Mardin and Siirt as places where *qeltu* dialects were still spoken.

Despite being a continuation of Mesopotamian dialects, Anatolian dialects of Arabic have been cut off from the mainstream of Arabic dialects. How exactly this cut-off and separation between dialects happened, given the lack of specific barriers, is largely unknown and remains at a speculative level. Regarding this topic, Procházka (this volume) suggests “the foundation of nation states after World War I entailed significant decrease in contact between the different dialect groups and an almost complete isolation of the Arabic dialects spoken in Turkey”.

Like Central Asian Arabic and Cypriot Maronite Arabic (see Walter, this volume), Anatolian Arabic dialects are characterized by (i) separation from the Arabic speaking world, (ii) contact with regional languages, which has affected them strongly, and (iii) multilingualism of speakers.

The Anatolian dialects have diverged much more from the Standard type of Arabic compared to the other *qeltu* dialects, e.g. Tigris or Euphrates groups (Jastrow 2011b). One of the hallmarks of Anatolian Arabic is the suffix *-n* instead of *-m* in the second and third person plural (e.g. in Mardin Arabic *baytkən* ‘your (PL) house’, *baytən* ‘their house’) and the negation *mō* with the imperfect. In addition to many interesting properties like the ones just mentioned, Anatolian Arabic

has acquired a large number of interesting contact-induced patterns.

These dialects are spoken as minority languages by speakers belonging to different ethnic or religious groups. As noted by Jastrow (2006), not all of the Anatolian Arabic varieties are spoken *in situ* however, and in fact some may no longer be spoken at all. Jastrow notes that some of the dialects were exclusively spoken by Christians and almost died out during World War I as a result of the massacres of the Armenians and other Christian groups. A few thousand speakers of these dialects survive to this day, most of whom have migrated to big cities, starting from the mid-1980s, particularly Istanbul. Some speakers of these dialects also live in Europe. Nevertheless, these dialects are very likely to face extinction in a few decades.

The Jews who spoke Anatolian Arabic varieties (mainly in Diyarbakır) migrated to Israel after the foundation of the State of Israel in 1948. These dialects also face a serious threat of extinction.

Today Anatolian Arabic dialects are predominantly spoken by Muslims (although there are a few hundred Arabic-speaking Christians, particularly in some parts of Istanbul, such as Samatya). These dialects are still found *in situ*, however they are also subject to constant linguistic pressure from Turkish (the official language) and Kurdish (the dominant regional Indo-Iranian language), and social pressure to assimilate. The quote from Grigore (2007b: 27) summarizes the overall context of Anatolian Arabic: “il se situe dans un microcontexte kurde, situé à son tour dans un macrocontexte turc, étant isolé de la sorte de la grande masse des dialectes arabes contemporains”.²

The total number of speakers lies around 140,000 (Versteegh 2014: 282), most of whom are bi- or trilingual in Arabic, Kurdish and Turkish. As Jastrow (2011a: 88) points out, the phenomenon of diglossia is not observed in Anatolia; instead Turkish occupies the position of the ‘High variety’, and Anatolian Arabic, the ‘Low variety’, occupies a purely dialectal position. In addition, speakers of different dialects may speak other minority languages as well. For instance, a considerable number of Sason Arabic speakers know the local variety of the Iranian language Zazaki, and those of Armenian origin speak an Armenian dialect.

Anatolian Arabic varieties are in decline among the speakers of these varieties, and public life is dominated primarily by Turkish (and Kurdish). The presence of Arabic in Turkey has increased due to Syrian refugees who fled to Turkey, yet this increased presence concerns Syrian Arabic, not Anatolian Arabic (see Procházka, this volume). In addition to the absence of awareness about Anatolian Arabic dia-

²“It is situated in a Kurdish microcontext, which is in turn situated in a Turkish macrocontext, thus being isolated from the vast majority of contemporary Arabic dialects”.

Faruk Akkuş

lects in the Arab states, the Anatolian dialects also suffer from a more general lack of interest. The speakers generally do not attribute any prestige to their languages, calling it “broken Arabic”, and often making little effort to pass it on to the next generations. It should, however, be noted that an increasing interest in these dialects has taken place in recent years, especially at the academic level. To this end, several workshops have been organized at universities in the relevant regions, aimed towards promoting these dialects and discussing possible strategies for their preservation.

The data referenced in this chapter come from various Anatolian Arabic dialects. The name of each variety and its source(s) are as follows: Āzəx (Wittrich 2001); Daragözü (Jastrow 1973); Hapəs (Talay 2007); Hasköy (Talay 2001; 2002); Kinderib (Jastrow 1978); Mardin (Jastrow 2006; Grigore 2007a; Grigore & Biţună 2012); Mutki-Sason (Akkuş 2016; 2017; Isaksson 2005); Siirt (Biţună 2016; Grigore & Biţună 2012); Tillo (Lahdo 2009).

2 Contact languages

2.1 Overview

Anatolia, especially the (south)eastern part, has been home to many distinct linguistic groups (as well as ethnic and religious groups). Up until the beginning of the twentieth century, speakers of the largest Anatolian languages – Kurdish, Zazaki, Armenian, Aramaic, Arabic and Turkish – had been co-existing for almost a thousand years. This has naturally resulted in extensive contact among these languages.

Contact influence on Anatolian Arabic has arisen mainly through long-term bi- and multi-lingualism rather than through language shift (in which speakers of other languages shifted to Arabic; Thomason 2001). As a result, when applicable, the changes seem to be primarily through borrowing, rather than imposition (in the sense of Van Coetsem 1988; 2000).

2.2 Turkish

Turkish, as the official language of Turkey, currently dominates public life in most Arabic-speaking areas. However, as noted by Haig (2014: 14), “the current omnipresent influence of Turkish in the region is in fact a relatively recent phenomenon, fueled by compulsory Turkish-language state education, the mass-media, and large-scale military operations carried out by the Turkish army in the

conflict against militant Kurdish groups. But prior to the twentieth century, the influence of Turkish in many parts of rural east Anatolia was negligible.”

Although Turkish is the dominant language in the public sphere, there are still many people, particularly in rural parts of the (south-)eastern Turkey, who do not speak Turkish, including speakers of Anatolian Arabic varieties. It is usually women over forty years old that fall into this category. They tend to speak the local Arabic variety along with the dominant language in that geographic area.

Moreover, the amount of Turkish influence is greater on the Arabic speakers who have migrated to bigger cities such as Istanbul, compared to those who still speak their dialects *in situ*.

2.3 Kurdish and Zazaki

Anatolian Arabic has been in intensive contact with two western Iranian languages: Kurmanji and Zazaki. These languages have influenced each other on different levels. As noted by Procházka (this volume) and Öpengin (this volume), Kurdish and Arabic, including the region of south-eastern Anatolia, have experienced extensive contact since at least the tenth century.

Due to the multi-ethnic (and to a lesser extent multi-religious) nature of the regions, bilingualism between Arabic and Kurdish (or Zazaki) is very widespread. The speakers of the non-dominant languages tend to have a stronger command of the dominant languages than the reverse situation. For instance, in Mutki, Bitlis province, where Kurdish is the dominant regional language, Arabic speakers have a native-like command of Kurdish, whereas not many Kurdish speakers speak the local Arabic variety. In some parts of Sason, Batman province, on the other hand, Arabic is the dominant language, and Kurdish speakers learn Arabic as a second language.

2.4 Aramaic

Aramaic and Arabic have for centuries lived side by side, so that it is possible to speak of both substrates (from Syriac/Neo-Aramaic to Arabic), and of adstrates, or rather, of superstrates (from Arabic to Aramaic). In the context of Anatolian Arabic, Aramaic has been in contact mainly with the Mardin dialect group.

These two languages have influenced each other in many ways. For instance, the many dialects constituting Modern Eastern Aramaic show considerable diversity as to choice of verbal particles. Some dialects use particles similar in form and function to those of the *qəltu*-dialects (see e.g. Jastrow 1985; as well as Coghill, this volume, for North-Eastern Neo-Aramaic dialects).

Faruk Akkuş

Finally, it is worth mentioning that, given the existence of Arabic speakers of Armenian origin, Armenian might have influenced certain Anatolian Arabic varieties. However, the influence of Armenian is hardly known, apart from the fact that many villages in the further eastern part of Anatolia, in which Arabic was spoken or is still spoken, bear Armenian names. This requires further investigation in its own right.

3 Contact-induced changes in Anatolian Arabic

Anatolian Arabic dialects manifest considerable variation, and have also come to exhibit interesting patterns due to language contact in every linguistic aspect. This section surveys these changes and features in turn.

3.1 Phonology

Anatolian Arabic has undergone significant changes in its consonant and vowel inventories due to language contact (as well as language internal developments). These changes include the introduction of new consonantal phonemes, loss or weakening of emphatic consonants, and introduction of new vowels. In addition to these changes, it is possible to count word-final devoicing as a contact-induced change.

This section first introduces the consonant inventory in Anatolian Arabic varieties. It should be noted that not all consonants are present in every variety, but the chart serves as the sum of consonants available across Anatolian Arabic varieties. For instance, the phonology of Sason Arabic (and other varieties of the Kozluk–Sason–Muş group) is characterized by the (near) absence of pharyngeal and emphatic (pharyngealized) consonants,³ which have fused with their plain counterparts, e.g. *pasal* ‘onions’ in Sason < Old Arabic (OA) *baṣal*.⁴

Table 1, with information largely taken from Jastrow (2011a), demonstrates that Anatolian Arabic has several consonants that were originally alien to Arabic (see 3.1.1 for discussion). With respect to the inventory of vowels, the noteworthy development is the introduction of /ē/ and /ō/ for some lexical items. Note that the Old Arabic diphthongs *ay and *aw have largely been preserved in these varieties:

³These sounds, whose emphatic quality is indicated in Table 1 and throughout with a subscript dot are only nearly absent for two reasons: (i) it is possible to detect them in the speech of elderly speakers in some lexical items, while the younger generations have lost them, (ii) Talay (2001) reports their availability in Hasköy, Muş province to a certain extent.

⁴Compare Cypriot Maronite Arabic (Walter, this volume), Maltese (Lucas & Čéplö, this volume) and Nigerian Arabic (Owens, this volume)

Table 1: Inventory of consonants. Marginal or doubtful phonemes within parentheses

	Labial	Interdental	Dental	Postalveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b		t ɾ d ɖ			k g	q		(ʔ)
Affricate				č ğɟ					
Fricative	f v	θ ð ɸ	s ʃ z	š ž		ħ ʁ		ħ ʕ	h
Nasal	m		n						
Vibrant			r ɾ						
Lateral			l						
Approximant	w				y				

Jastrow (2011a: 89) notes that one of the processes by means of which these mid long vowels entered the inventory of Anatolian Arabic is via loanwords from Turkish and Kurdish, e.g. commonly used items, *čöl* ‘desert’, *təl* ‘wire’ (Turkish, probably through the intermediary of Kurdish), *ħōrt* ‘young man’ (Kurdish).

3.1.1 New phonemes ‘p, č, ž, g, v’

The Anatolian Arabic varieties, as well as the varieties in (northern) Syria and Iraq, have certain phonemes that were not originally familiar to these varieties of Arabic. These phonemes include the voiceless bilabial stop /p/, the voiceless affricate /č/, the voiced post-alveolar fricative /ž/,⁵ the voiced velar stop /g/, and the voiced labiodental fricative /v/.⁶ The emergence of these phonemes is most likely due to the massive contact with Turkish, Kurdish and Aramaic. That is, the most likely scenario is that the centuries-long borrowing of words which contained these sounds ultimately resulted in them getting incorporated into the phonemic inventory.

⁵Cf. Jastrow (2011a) and Grigore & Bițună (2012) regarding the status of /ž/: this sound is largely restricted to borrowed words. The reflex of Arabic <ع> in Anatolian Arabic is /ğ/.

⁶Blanc (1964: 6–7) considers /p/ and /č/ as characteristic of Mesopotamian varieties.

With regard to /v/, it is likely that there are two paths of emergence: (i) as an internal evolution of the voiced interdental fricative /ð/ and (ii) via loan-words from Turkish and Kurdish. The forms *vîp* and *zîp* ‘wolf’ (cf. OA *ðiʔb* ‘wolf’) represent a language internal development, whereby the interdental fricatives have shifted to sibilants in Kozluk-Sason-Muş, and to labiodental fricatives in Āzəx (Şırnak province, Wittrich 2001), whereas they have been retained in most Mardin group dialects.⁷

In many cases, it is impossible to pinpoint which language these sounds were (initially) borrowed from. However, as also noted in Procházka (this volume), /p/ was probably introduced via contact with Kurdish, followed by influence from Ottoman and modern Turkish.⁸ Some illustrations are as follows:

- (1) *pîs* ‘dirty’, cf. Kurdish/Turkish *pîs*, *pis*
parčāye ‘piece’, cf. Turkish *parça*
pûz ‘nose’ (Hapəs), cf. Kurdish *pûz*
davare ‘ramp’, cf. Kurdish *dever* fem. ‘place’
čuvāle ‘sack’, cf. Turkish *çuval*
pêlāv (Hasköy) ‘shoe’, cf. Kurdish *pêlav*
čāy ‘tea’, cf. Turkish *çay*
čaqmāq ‘lighter’, cf. Turkish *çakmak*
rēnčbarî (Hasköy), *rēžbarî* (Sason) ‘husbandry’, cf. Kurdish *rênjbarî*
žîžo (Āzəḥ) ‘hedgheg’, cf. Kurdish *žîžo*
tāži ‘greyhound’, cf. Kurdish *tajî*
gōmlak ‘shirt’, cf. Turkish *gömlek*
magzūn, *mazgūn* (in Sason) ‘sickle’, cf. Syriac *magzūnā*; Turoyo *magzūno*

Talay (2007) suggests that the loss of the phonemic status of the emphatic consonants and the weakness of the pharyngeal in Kozluk–Sason–Muş group is likely due to the influence of Turkish, which does not have them. Examples are from the Hasköy dialect, and are taken from Talay (2007: 181):

- (2) *ata* ‘he gave’ (< *ʔaʔtā), cf. *adā* in Sason
sēbi ‘boy’ (< *šabiyy)
zarab ‘he hit’ (< *ḏarab < *ḏarab)

Thus, changes of this kind can be seen as a quasi-adaptation of the consonant inventory to that of the superstrate languages.

⁷For more discussion, see Wittrich (2001), Jastrow (2011a), Grigore (2007a), Talay (2011), Akkuş (2017), and Biṭunā (2016) among others.

⁸For further illustrations and discussion, see e.g. Vocke & Waldner (1982), Jastrow (2011a), Talay (2002; 2007) and Grigore & Biṭunā (2012).

3.1.2 Word-final devoicing

Certain voiced stops in Anatolian Arabic /b, d, ġ, g/ have a tendency to become devoiced [p, t, tʃ, k] when they occur word-finally, probably due to Turkish influence, which is well-known for this property.

For instance, /b/ is mainly realized as the voiceless [p] in final pre-pausal position, e.g.: *anep* ‘grape(s)’, cf. OA *ʃinab*; *yarīp* ‘stranger’, cf. OA *yarīb*. This might reflect a change in progress, as [Lahdo \(2009\)](#) points out that the incidence of devoicing in other Anatolian dialects is also increasing over time. Note that the devoicing process does not take place in all instances, supporting the claim that the language is undergoing a transition in this regard. Moreover, the lack of a written form removes a possible brake on this process. Further illustrations are as follows:

- (3) *axa*[θ] ‘he took’
kata[p] ‘he wrote’ (Mardin; [Jastrow 2011a](#): 90)
ktē[p] ‘book’ (Mardin), cf. OA *kitāb*
baʃi[t] ‘far’ (Āzəx), cf. OA *baʃid*
aṭya[p] ‘nicer’ (Tillo), cf. OA *ʔaṭyab*
azyā[t] ‘more’, cf. OA *ʔazyad* ([Lahdo 2009](#): 106)

Devoicing is not limited to word-final position, however, but is also attested before voiceless consonants, e.g. *haps* ‘prison’, cf. OA *ḥabs*.

3.2 Morphology

3.2.1 Reduplication

A type of reduplication due to contact with Turkish produces doublets with /m/. The consonant /m/ may be added initially to vowel-initial words, as in (4a), or replaces the initial consonants in consonant-initial words, as in (4b) (see [Akkuş 2017](#); [Lahdo 2009](#)). The new meaning is either that of vagueness or ‘*et cetera*’.

- (4) Sason Arabic
 a. *ažin* *m-ažin*
 dough m-dough
 ‘dough or something like that’
 b. *hās* *m-ās* *lā* *təso*
 sound m-sound NEG make.IMPF.2PL
 ‘Don’t make any noise!’ (Lit. ‘Don’t make sound or something like that.’)

Faruk Akkuş

Following the same restriction in Turkish, if a word starts with /m/, this type of reduplication is disallowed, e.g. *māse* ‘table’ cannot be reduplicated in a way that would result in *māse māse*.

3.2.2 Degree in Adjectives

Adjectives in Anatolian Arabic follow the noun directly, agreeing with it in gender, number, and definiteness. In this respect, the situation is similar to most Arabic varieties. Degree, on the other hand, is not an inflectional category in Sason Arabic. Instead, this dialect has adopted the Turkish adverbs *daha* ‘more’ and *en* ‘most’ for comparative and superlative, respectively. Both these items precede the adjectival constituent, as shown in (5a) and (5b).

(5) Sason Arabic

- a. mən-n-i daha koyse ye
 from-OBL.1SG more beautiful.F COP.3SG
 ‘She is more beautiful than me.’
- b. en gbīr
 most big
 ‘the biggest’

The Tillo variety also uses the Turkish-derived *an* ‘most’ in superlative forms, with both Arabic-derived adjectives (in the elative form) and Turkish-derived adjectives (which lack an elative form), as in (6a) and (6a) respectively.

(6) Tillo Arabic (Lahdo 2009: 198)

- a. an⁹ at̤yap
 most delicious.ELAT
 ‘the most delicious’
- b. an yāqən
 most close
 ‘the closest’

On the other hand, the comparative in the Tillo variety is formed through the elative alone (which functions in other Arabic varieties as both comparative and superlative). The standard of comparison is introduced by the preposition *mən* ‘from’.

⁹Lahdo (2009) describes this vowel as “short front-to-back unrounded” in Tillo.

- (7) Tillo Arabic (Lahdo 2009: 162)
 təllo iyy aṭyap mən əṣṭanbūl
 Tillo COP.3SG good.ELAT than Istanbul
 ‘Tillo is better than Istanbul.’

3.2.3 Derivational affixes

Through numerous loanwords, certain derivational suffixes have been introduced into Anatolian Arabic. These suffixes include the agentive morpheme *-ğī/-čī*, and the abessive suffix *-səz*, which translates as ‘without’. Ingham (2011: 178) points out that these suffixes, especially the former, are also found in the dialects of Iraq, Syria, and elsewhere.

- (8) a. Sason Arabic
 gahwa-ğī
 coffee-AGT
 ‘coffee maker’
 b. Sason Arabic
 viğdan-səz
 conscience-ABESS
 ‘unconscientious’
 c. Tillo Arabic (Lahdo 2009: 199)
 kəlla kānu mṭahhər-čī-yye
 all be.PRF.3PL circumcizer-AGT-PL
 ‘They all were circumcizers.’

The presence of these suffixes on lexemes of the local Arabic varieties, e.g. *ḥāser-ğī* ‘yogurt maker, yogurt seller’ (Sason Arabic) or *mṭahhər-čī* ‘circumsiser’ (Tillo Arabic), suggests that the forms above are not necessarily adopted as a whole. Rather, Arabic speakers may decompose the word and apply the suffix to other lexemes in some cases.

3.2.4 Compounds

Anatolian Arabic has borrowed the N+N compounding strategy from Turkish, where the right-hand member carries the compound linker morpheme *-i*. This pattern is not generally found in other varieties of Arabic and it is most likely due to contact with Turkish. This type of compound is often used with whole Turkish phrases. The examples are as follows (note that the buffer consonant *-s*

Faruk Akkuş

appears between the linker morpheme and the noun when the noun ends in a vowel):

- (9) Sason Arabic (Akkuş & Benmamoun 2018: 41)

- a. lisa mudur-i
 high.school director-LINK
 ‘high school director’
- b. qurs oratman-i
 course teacher-LINK
 ‘course teacher’

- (10) Tillo Arabic (Lahdo 2009: 199)

- fəstəq fabriqa-si
 pistachio factory-LINK
 ‘pistachio factory’

This compounding strategy is found in other Arabic varieties spoken in Turkey as well, for instance, *buz dolab-i* ‘refrigerator’ (lit. ‘ice cupboard-LINK’) in Adana dialect. Whether compounding has been borrowed as a productive process as opposed to borrowing of the whole phrase requires further investigation.¹⁰

3.2.5 Vocative ending -o

Another morphological feature that Anatolian Arabic has acquired is the vocative particle -o. When addressing a person directly, -o is commonly affixed to kinship terms and given names. This appears to be available in the whole area. Unlike the situation in Syria and Iraq (see Procházka, this volume), this form of address is not usually used hypocoristically. Some examples are below:

- (11) *amm-o* ‘(paternal) uncle!’
 ġemāl-o ‘Cemal!’
 ḥāl-o ‘(maternal) uncle!’

The corresponding forms of feminine nouns end in -ē, as in *ḥabībt-ē* ‘darling!’. Grigore (2007b: 203) suggests that this vocative -o is borrowing of a morphological form from Kurdish (cf. Haig & Öpengin 2018), since the suffix, with masculine and feminine forms, is not historically available in Arabic.

¹⁰Thanks to Stephan Procházka for the discussion and the example from the Adana dialect.

In brief, contact with Turkish and other neighboring languages has led to various noticeable changes in the morphology of Anatolian Arabic, particularly the more easterly varieties.

3.3 Syntax

Research on the syntax of Anatolian Arabic varieties, let alone work on contact-induced syntactic changes, significantly lags behind the research conducted on other aspects of these languages. Several factors might have contributed to this effect. The researchers' tendency to focus on phonological or lexical aspects and the lack of sufficient data from which to draw conclusions are two possible factors. Another possibility that Ingham (2005) raises for contact-induced syntactic change is that since the languages in contact are so typologically different, it is difficult for them to adopt syntactic features from each other without extensive language change taking place.

This section introduces several syntactic phenomena that can be attributed to language contact, including copulas, marking of indefiniteness, light verb constructions and the periphrastic causative. Although the details are not elaborated on here, the conclusion we can arrive at is in line with Ingham (2005), in that the degree and intensity of contact with the neighboring languages leads to differences among Anatolian Arabic dialects. The more easterly varieties, e.g. the Kozluk–Sason–Muş group, appear to be the most innovative, and the dialect group(s) most influenced by the language contact, whereas the Mardin group appears to be the most conservative (see Akkuş 2017; Jastrow 2011a for more discussion).

3.3.1 Copula

One of the most distinctive features of Anatolian Arabic is the existence of the copula in nominal sentences, based on the independent pronouns. This copula is realized as an enclitic suffix in most Anatolian dialects. Although researchers seem to differ with respect to the degree of the influence, they converge on the view that it is a matter of language contact, and that at least the development and the proliferation of the obligatory copula is under the influence of the neighboring languages – Turkish, Kurdish, Zazaki and Aramaic – which all have copulas in nonverbal clauses (see Lahdo 2009; Grigore 2007a; Palva 2011; Talay 2007; Jastrow 2011a; Akkuş 2016; 2017; Akkuş & Benmamoun 2018, for more discussion and illustrations).

Faruk Akkuş

Although the copula forms themselves are not imported, the way they are used in Anatolian Arabic is exactly the same as it is in Kurdish, Turkish and ʿTuroyo (Aramaic), which have copula in the present tense. The copula is placed after the predicate (examples from Grigore 2007a).

- (12) a. Kurdish
bave-min şivan-e
father-POSS.1SG shepherd-3SG
'My father is a shepherd.'
- b. Turkish
baba-m çoban-dır
father-POSS.1SG shepherd-3SG
'My father is a shepherd.'
- c. ʿTuroyo
bab-i rəŋyo-yo
father-POSS.1SG herder-3SG
'My father is a herder.'

Some examples from Anatolian Arabic are illustrated below:¹¹

- (13) a. Kinderib Arabic (Jastrow 1978: 131)
malih-we
beautiful-3SG.M
'He is beautiful.'
- b. Sason Arabic
rahw-în nen
sick-PL 3PL
'They are sick.'
- c. Daragözü Arabic (Jastrow 1973: 40)
nā ḫāš-nā
1SG good-1SG
'I am good.'

¹¹It should be noted that the copula is not necessarily realized as an enclitic in some dialects. For instance, in the dialect of Siirt (Jastrow 2011a) the copula precedes the predicate. Moreover, the copula is identical to the personal pronoun in Siirt, whereas other Anatolian varieties use the shortened version of the pronoun in the 3SG and 3PL. See Akkuş (2016) for some discussion.

In negative sentences as well, the same order of morphemes is attested. The negative morpheme (and the copula if there is one) follows the predicate in the neighboring languages, as the sentences in (14) show.

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

The same order is found in Sason Arabic, in that the NEG+COP follows the predicate.¹²

- (15) Sason Arabic
 nihane me-nnen
 here NEG-COP.3PL
 ‘They are not here.’¹³

Given that the copula is almost unknown in other Arabic speaking areas (but see [Blanc 1964](#); also [Lucas & Čéplö](#), this volume; [Walter](#), this volume), it is safe to assume that the development of a full morphological paradigm for the copula along with its syntactic function is at least facilitated by contact with the neighboring languages.

3.3.2 Light verb construction

Light verb constructions are another domain where the influence of contact is clearly manifested. In surrounding languages, particularly Turkish, Kurdish and

¹²This is not the most common order in Anatolian Arabic varieties, however. For more discussion, see [Jastrow \(2011a\)](#) and [Akkuş \(2016; 2017\)](#).

¹³In Sason Arabic, the 3PL personal pronoun can be *innen* or *iyen*. A shortened version of this pronoun is used both in affirmative, as in (13b) and negative, as in (15), non-verbal clauses.

Faruk Akkuş

Zazaki, a light verb construction consists of a nominal part followed by the light verb, which is usually ‘to do’ or ‘to be’, e.g. Kurdish *pacî kirin* (lit. ‘kiss do’) ‘to kiss’, Turkish *motive etmek* (lit. ‘motivation do’) ‘to motivate’.

There are a relatively large number of compound verbs constructed with Arabic *sāwa* – *ysāwi* ‘to do’ and a nominal borrowed from Turkish or Kurdish, as illustrated in (16). In the majority of the cases, the construction is a complete calque of its Turkish or Kurdish counterparts (see e.g. Versteegh 1997; Lahdo 2009; Grigore 2007a; Talay 2007; Jastrow 2011a; Akkuş 2016; 2017; Akkuş & Benmamoun 2018; Biṭună 2016 for more examples).

- (16) a. Āzəḥ/Mardin Arabic (Talay 2007: 184)
sāwa brîndār ‘to injure’ cf. Kurdish *brîndar kirin*
sāwa ḡāmêrtîye ‘to act generously’ cf. Kurdish *camêrtî kirin*
sāwa yôt ‘to mow’ cf. Kurdish *ḡôt kirin*
- b. Tillo Arabic (Lahdo 2009: 202)
sāwa yārdəm ‘to help’ cf. Turkish *yardım etmek*
ysāwaw dawām ‘they continue ...’ cf. Turkish *devam etmek*
nsayy qaḥwaltə ‘we have breakfast’ cf. Turkish *kahvaltı etmek*

In Sason Arabic, the default order in this construction has reversed, in that in most cases the nominal is followed by the light verb. Thus, Sason manifests head-final order, undoubtedly due to contact with Turkish and Kurdish. Similarly, the nominal part of the construction can be borrowed from Turkish as in (17), including instances of reborrowing of an originally Arabic word, (17b), or Kurdish as in (18). In fact, the nominal part might also be Arabic, as in (19).

- (17) Sason Arabic (Turkish borrowing)
- a. qazan *sāwa*
win do.PRF.3SG.M
‘to win’
- b. iṣāret *sāwa*
sign do.PRF.3SG.M
‘to sign’
- (18) Sason Arabic (Kurdish borrowing)
ser asi
watch do.IMPF.1SG-do
‘I watch ...’
- (19) Sason Arabic

2 Anatolian Arabic

- a. gerre/hās sāwa
noise/sound do.PRF.3SG.M
'to make noise/sound'
- b. šəyle lā təsi, aməl si!
talk NEG do.IMPF.2SG.M work do.IMP.2SG.M
'Don't talk, do work!'
- c. huğūm sinna
attack do.PRF.1PL
'We attacked.'

Anatolian Arabic usually resorts to the same periphrastic construction when borrowing verbs from Turkish; it creates a complex predicate, rather than adapting a foreign verb directly to Arabic verbal morphology, a borrowing strategy seen also in the other languages in the region, such as Kurdish, Zazaki. In many cases, the complex predicate comprises of *sāwa* + the Turkish verbal form of the indefinite past (i.e. *miş*-verb), rather than the bare form of the verb, as illustrated in (20).

- (20) Anatolian Arabic (Talay 2007: 184)
sawa gačənməš 'to manage'; cf. Turkish *geçinmiş*
başlaməš sawa 'to begin'; cf. Turkish *başlamış*

Despite the widespread use of this process for loanwords, some borrowed verbal forms have been totally assimilated to the Arabic verbal system; the majority of these verbs are formed according to verbal measures (stems) II or III, as can be seen in example (21).

- (21) Āzəḥ (examples from Talay 2007)
- | | | |
|----------|---|---------------------------------|
| Stem II | <i>qappat – iqappət</i> 'to close' | cf. Turkish <i>kapatmak</i> |
| Stem II | <i>qayyad – iqayyəd</i> 'to register' | cf. Turkish <i>kayıt etmek</i> |
| Stem III | <i>dāyan – idāyən</i> 'to be patient, to bear up' | cf. Turkish <i>dayanmak</i> |
| Stem III | <i>tēlan – itēlən</i> 'to rob' | cf. Kurdish <i>talani kirin</i> |

3.3.3 Marking of (in)definiteness

In Classical Arabic and in modern varieties spoken in the Arab world, the indefinite noun phrase is unmarked or is preceded by an independent indefinite particle, whereas an NP becomes definite by prefixing the definite article *al-/əl-/l-* etc. (Brustad 2000). However, Kozluk–Sason–Muş group dialects have adopted

Faruk Akkuş

the reverse pattern (see also Uzbekistan Arabic; [Jastrow 2005](#)), which is found in the neighboring languages Turkish and Kurdish. That is, the definite NP is left unmarked, and the enclitic *-ma* is used to mark the indefiniteness of an NP ([Talay 2007](#); [Akkuş 2016](#); 2017; [Akin et al. 2017](#); [Akkuş & Benmamoun 2018](#)), as illustrated in (22).

- (22) Sason Arabic
mara ‘the woman’ > *mara-ma* ‘a woman’
bayt ‘the house’ > *bayt-ma* ‘a house’

The parallel constructions in Kurdish and Turkish are illustrated in (23) and (24) respectively.

- (23) Kurdish
derî ‘the door’ > *derî-yek* ‘a door’
 (24) Turkish
kadın ‘the woman’ > *bir kadın* ‘a woman’ (Turkish)

3.3.4 Periphrastic causative

Sason Arabic resorts to periphrastic causative 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 *bîdîn* ‘to give’ to form the causative, as in (25).

- (25) Kurdish ([Atlamaz 2012](#): 62)
 mî piskilet do çekır-in-e
 OBL.1SG bicycle give.PTCP repair.PTCP-GER-OBL
 ‘I had the bicycle repaired.’ (Lit: ‘I gave the bicycle to repairing.’)

Sason Arabic exhibits the same pattern for causative and applicative formation, as shown in (26), which is most likely as a result of extensive contact with Kurdish.¹⁴

¹⁴Sason Arabic also has another periphrastic construction that is formed with the verb *sa* ‘to do/make’, which may embed a finite clause (i.a) or a verbal-noun phrase (i.b).

- (i) Sason Arabic (adapted from [Taylan 2017](#): 221)
 a. dohtor masha ali ku isi fiy-u (le yaddel)
 doctor to Ali COP.3SG.M make.IMPF.3SG.M in-3SG.M (COMP make.IMPF.3SG.M)

- (26) Kurdish (Taylan 2017: 221)

əmm-a məša fatma ši adəd-u addil
 mother-OBL.3SG.F to Fatma food give.PRF.3SG.F-3SG.M make
 ‘Her mother made Fatma cook (Lit: Her mother gave food making to Fatma).’

3.4 Lexicon

Anatolian Arabic dialects have borrowed single words and whole phrases or expressions mainly from (Ottoman and modern) Turkish and Kurdish. The influence of these two languages on the Arabic lexicon is enormous. Aramaic words also survive in Anatolian Arabic to a lesser degree. A few illustrations are given in (27).

- (27) *bōš* ‘much’, cf. Kurdish *boş*
bōšqa ‘different’, cf. Turkish *başka*
ṛūvi ‘fox’, cf. Kurdish *rûvû*
hič ‘none, whatsoever’, cf. Turkish *hiç*
səpor ‘sport’, cf. Turkish *spor*
magzūn, *mazgūn* (in Sason) ‘sickle’, cf. Syriac *magzūnā*; Țuroyo *magzūno*

As Jastrow (2011a: 95) mentions, while more Turkish borrowings are found in bigger cities such as Mardin, Diyarbakır or Siirt, Kurdish borrowings constitute a bigger part of the lexicon of rural dialects. Anatolian Arabic dialects which have preserved the emphatics, pharyngeals or interdental adapt borrowings into their phonology. For instance, Turkish *halbuki* ‘however’ is borrowed as *ḥālbūki*. In most cases, the velar *k* is turned into the uvular *q*, e.g. *čaqmāq* ‘lighter’, cf. Turkish *çakmak*. Also, Kurdish feminine nouns (and even some Turkish nouns) are suffixed with the Arabic feminine morpheme *-e/-a*, e.g. *tūre* ‘shoulder’ (cf. Kurdish *tūr*).

sipor
 sports

‘The doctor is making Ali do sports.’

- b. aya sa hazd hašiš
 landlord make.PRF.3SG.M cut.INF grass

‘The landlord had the grass cut.’

Although the origin of these constructions is not clear, they do not appear to be a contact-induced.

Faruk Akkuş

There are several function words that are copied from Turkish into Arabic, e.g. Turkish *ama* ‘but’ is realized as *hama* in Sason, and as *aṃa* in Tillo Arabic. The conjunction *çünkü* ‘because’ from Turkish is attested in many Anatolian varieties, with the same function. Lahdo (2009: 179) notes that it expresses causal clauses in Tillo, as in (28), and Biṭunā (2016: 213) states the same role for Siirt. Jastrow (1981: 278) and Grigore (2007b: 261) also confirm its existence in Ḥalanze and Mardin, respectively.

(28) Tillo Arabic (Lahdo 2009: 179)

mā ‘aṭaw-ni əzan čünki ğītu əl-anqara
NEG give.PRF.3PL-1SG permission because come.PRF.1SG DEF-Ankara
‘They did not give me permission because I had come to Ankara.’

Procházka (2005) notes that particles such as *bīle* < *bile* ‘even’, or *zātan* < *zaten* ‘already’ in the Adana region are also borrowed from Turkish (see also Isaksson 2005).

4 Conclusion

This chapter has dealt with contact-induced changes in the Anatolian Arabic dialects. We have seen that Anatolian Arabic has been primarily in contact with Turkish, Kurdish and Aramaic, and the influence of these neighboring languages on Anatolian Arabic is evident. We have surveyed some contact-induced changes at the phonological, morphological, syntactic and lexical level.

Mardin and Siirt dialects have been covered much more comprehensively than other dialects in the literature. It is desirable to have more comprehensive investigations carried out for the dialects around the Bitlis, Muş and Diyarbakır areas. This research has the potential to fill the gaps in our current state of knowledge about these dialects.

Similarly, in the literature, in terms of the linguistic features investigated, phonological and morphological properties (along with lexicon) have received more attention, whereas syntax, in particular, has been understudied. This situation, however, might change once we are at a point where we have enough recordings and transcriptions to investigate syntactic properties of the dialects.

Further Reading

Jastrow (1978) is a seminal work, which provided the classification for Anatolian Arabic varieties.

Jastrow (2011a) is a concise, yet comprehensive encyclopedia entry on characteristic features of Anatolian Arabic.

Talay (2011) is a good source for an overview of Arabic dialects in the Mesopotamian region.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	GER	gerund
I, II, III, ...	1st, 2nd, 3rd verbal derivation	IMPF	imperfect
		INF	infinitive
ABESS	abessive	LINK	linker
AGT	agentive	M	masculine
COMP	complementizer	NEG	negation
COP	copula	OA	Old Arabic
CyA	Cypriot Maronite Arabic	OBL	oblique
		PL	plural
DEF	definite article	POSS	possessive
ELAT	elative	PRF	perfect
F	feminine	PTCP	participle
		SG	singular

Acknowledgements

I would like to thank Stephan Procházka and Mary Ann Walter for sharing their work, and to Stephan Procházka for reading another version of the paper, which also helped with this version. Thanks to Gabriel Bițuňa for directing my attention to several important sources. I also thank the editors Chris Lucas and Stefano Manfredi for their help and patience. All remaining errors are of course my own.

References

Akın, Cahit, Otto Jastrow & Shabo Talay. 2017. Nörşen: Ein anatolisch-arabischer Dialekt der Sason–Muş–Gruppe. *Zeitschrift für Arabische Linguistik* 65. 73–96.

Akkuş, Faruk. 2016. The Arabic dialect of Mutki–Sason areas. In George Grigore & Gabriel Bițuňa (eds.), *Arabic varieties: far and wide. proceedings of the 11th international conference of aida, bucharest 2015*, 29–41. Bucharest: Editura Universității din București.

Faruk Akkuş

- Akkuş, Faruk. 2017. Peripheral Arabic dialects. In Elabbas Benmamoun & Reem Bassiouney (eds.), *The Routledge handbook of Arabic Linguistics*, 454–471. London: Routledge.
- Akkuş, Faruk & Elabbas Benmamoun. 2018. Syntactic outcomes of contact in Sason Arabic. In Manfredi Stefano & Mauro Tosco (eds.), *Arabic in contact*, 38–52. Amsterdam: John Benjamins.
- Atlamaz, Ümit. 2012. *Ergative as accusative case: Evidence from Adıyaman Kurmanji*. Istanbul: Bogaziçi University Master's dissertation.
- Biţună, Gabriel. 2016. *Morfo-sintaxa dialectului Arab Nord-Mesopotamian din Siirt, Turcia*. Bucharest: University of Bucharest doctoral dissertation.
- Blanc, Haim. 1964. *Communal dialects in baghdad*. Cambridge, MA: Harvard University Press.
- Brustad, Kristen. 2000. *The syntax of spoken Arabic: a comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Georgetown: Georgetown University Press.
- Grigore, George. 2007a. L'énoncé non verbal dans l'arabe parlé à Mardin. *Romano-Arabica*. 51–63.
- Grigore, George. 2007b. *L'arabe parlé à Mardin: Monographie d'un parler arabe "périphérique"*. Bucharest: Editura Universităţii din Bucureşti.
- Grigore, George & Gabriel Biţună. 2012. Common features of North Mesopotamian Arabic dialects spoken in Turkey (Şırnak, Mardin, Siirt). *Bilim, Düşünce ve Sanatta Cizre (Uluslararası Bilim Düşünce ve Sanatta Cizre Sempozyumu Bildirileri)*. 545–555.
- Haig, Geoffrey & Ergin Öpengin. 2018. Kurmanji Kurdish in Turkey: Structure, varieties and status. In Christiane Bulut (ed.), *Linguistic Minorities in Turkey and Turkic-speaking minorities of the peripheries*, 157–230. Wiesbaden: Harrassowitz.
- Haig, Geoffrey. 2014. East Anatolia as a linguistic area? Conceptual and empirical issues. In Lale Behzadi, Patrick Franke, Geoffrey Haig, Christoph Herzog, Birgitt Hofmann, Lorenz Korn & Susanne Talabardon (eds.), *Bamberger Orientstudien*, 13–36. Bamberg: University of Bamberg Press.
- Ingham, Bruce. 2005. Persian and Turkish loans in the Arabic dialects of North Eastern arabia. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: case studies from iranian, semitic and turkic*, 173–180. London: Routledge.
- Ingham, Bruce. 2011. Afghanistan Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

2 Anatolian Arabic

- Isaksson, Bo. 2005. New Linguistic Data from the Sason Area in Anatolia. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal divergence: case studies from Iranian, Semitic and Turkic*, 181–190. London: Routledge.
- Jastrow, Otto. 1973. *Daragözü: Eine arabische Mundart der Kozluk-Sason-Gruppe (Südostanatolien)*. Nürnberg: Hans Carl.
- Jastrow, Otto. 1978. *Die mesopotamisch-arabischen qeltu-Dialekte*. Vol. 1: *Phonologie und Morphologie*. Wiesbaden: Steiner.
- Jastrow, Otto. 1981. *Die mesopotamisch-arabischen qeltu-Dialekte. Band II, Volksskundliche Texte in elf Dialekten*. Wiesbaden: Steiner.
- Jastrow, Otto. 1985. *Laut- und Formenlehre des neuaramäischen Dialekts von Mîdin in Tûr ‘Abdîn*. Wiesbaden: Harrassowitz.
- Jastrow, Otto. 2005. Uzbekistan Arabic: A language created by Semitic–Iranian–Turkic linguistic convergence. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*, 133–139. London: Routledge.
- Jastrow, Otto. 2006. Arabic dialects in Turkey: Towards a comparative typology. *Türk Dilleri Araştırmaları* 16. 153–164.
- Jastrow, Otto. 2011a. Anatolian Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Jastrow, Otto. 2011b. Iraq: Arabic dialects. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Lahdo, Ablahad. 2009. *The Arabic dialect of Tillo in the region of Siirt (south-eastern Turkey)*. *Acta Universitatis Upsaliensis*. Uppsala: Uppsala University.
- Palva, Heikki. 2011. Dialects: Classification. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Procházka, Stephan. 2005. The Turkish contribution to the arabic lexicon. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*, 191–202. London: Routledge.
- Talay, Shabo. 2001. Der arabische Dialekt von Hasköy (Dêr Khâs), Ostanatolien, I: Grammatische Skizze. *Zeitschrift für Arabische Linguistik* 40. 71–89.
- Talay, Shabo. 2002. Der arabische Dialekt von Hasköy (Dêr Khâs), Ostanatolien, II: Texte und Glossar. *Zeitschrift für Arabische Linguistik* 41. 46–86.
- Talay, Shabo. 2007. The influence of Turkish, Kurdish and other neighbouring languages on Anatolian Arabic. *Romano-Arabica*. 179–189.

Faruk Akkuş

- Talay, Shabo. 2011. Arabic dialects of mesopotamia. In Stefan Weninger, Janet C. E. Watson, Michael Streck & Geoffrey Khan (eds.), *The Semitic languages: An international handbook*, 909–919. Berlin: De Gruyter Mouton.
- Taylan, Eser. 2017. Language contact in anatolia: the case of Sason Arabic. In Ramazan Korkmaz & Doğan Gürkan (eds.), *Endangered languages of the caucasus and beyond*, 209–225. Leiden: Brill.
- Thomason, Sarah G. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Versteegh, Kees. 1997. *The Arabic language*. First edition. Edinburgh: Edinburgh University Press.
- Versteegh, Kees. 2014. *The Arabic language*. Second edition. Edinburgh: Edinburgh University Press.
- Vocke, Sybille & Wolfram Waldner. 1982. *Der Wortschatz des anatolischen Arabisch*. Erlangen: W. Waldner.
- Wittrich, Michaela. 2001. *Der arabische Dialekt von Āzax*. Wiesbaden: Harrassowitz.

Chapter 3

Cypriot Maronite Arabic

Mary Ann Walter

Middle East Technical University, Northern Cyprus Campus

Cypriot Maronite Arabic is a severely endangered variety that has been in intensive language contact with Greek for approximately a millennium. It presents an interesting case of a language with extensive contact effects which are largely limited to the phonological domain.

1 Current state and historical development of Cypriot Maronite Arabic

Cypriot Maronite Arabic (CyA) is a minority language spoken by a small community on the island of Cyprus. Although essentially moribund, it is currently the focus of preservation and revitalization efforts.

1.1 Historical development of Cypriot Maronite Arabic

The time of arrival of this community of Arabic speakers to Cyprus is unknown. The island was occupied by an Arab garrison subsequent to Muḡāwīya's invasion of 649 CE, but the garrison was then removed and, presumably, the Arabic speakers left as well. More likely, a permanent presence dates back to the population movements of the ninth and tenth centuries during disruptions to Byzantine rule.¹ Subsequent waves of Arab emigration to Cyprus are documented during the early crusading period. Such movements were also quite likely during Lusignan (French crusader) rule in Cyprus (1192–1489), for some portion of which the Anatolian city of Adana, where Arabic is still widely spoken (see Procházka,

¹See §2 for a discussion of where the CyA-speaking community originated from and the dialectological affiliation of this variety of Arabic.



Mary Ann Walter

this volume), was also held by Lusignan rulers. Speakers of not only Arabic but a locally distinct version of Arabic in Cyprus are mentioned by Arab historians beginning in the thirteenth century, thereby providing a *terminus ad quem* to its dialectal development (Borg 2004).

As fellow communicants in the Catholic church, the Maronite community was granted certain privileges of independent worship during the Lusignan period, which were later lost during Venetian (1489–1571) and Ottoman rule (1571–1878), at which time some retaliation occurred on the part of the Orthodox community (Gulle 2016). After the Ottoman conquest of Cyprus in 1571, the Maronite community was at first placed under the administration of the Orthodox bishop, but regained religious autonomy shortly thereafter.

The number of Maronite villages underwent a steady decline during the Ottoman period, from over thirty to only five at the time of British occupation of the island in 1878 (Baider & Kariolemou 2015; though it is unclear if this is associated with any actual population decline). The remaining five villages are all located in the northwestern area of the island. However, as of the twentieth century at least, only one of them was home to speakers of CyA, the others having linguistically assimilated to Cypriot Greek entirely. The CyA-speaking village is Kormakiti(s) (also known as Kormacit and Koruçam in CyA and in Turkish, respectively).

Both the Cypriot liberation struggle of the 1950s against the British, and the years after independence was attained in 1960, saw increased communal conflict between the Turkish and Greek communities on the island. This period witnessed increasing separation of communities, as Turkish Cypriots withdrew into ethnic enclaves, and culminated in the 1974 conflict between Greece-sponsored coup plotters, military forces of Turkey, and local Cypriots on various sides, the result of which was a *de facto* division of the island between the Republic of Cyprus-controlled territory in the south, which was majority Greek Orthodox and Greek-speaking, and the Muslim and Turkish-speaking northern part of the island. This northern area subsequently declared independence, but remains unrecognized by any other country except the Republic of Turkey to this day.

It is important to note that the relative geographical separation between Greek Cypriots and Turkish Cypriots dates only from this recent period, as refugees sought safety within their own communities. This entailed a radical change in the social circumstances of CyA speakers, who moved to the capital city of Nicosia essentially *en masse*. Thus, they went from living in a Maronite village in which community life could be conducted in CyA, to being a tiny percentage of a large urban population. Not only that, but the pre-1974 population surrounding the CyA-speaking Maronite village of Kormakiti was composed of Greek speakers,

3 *Cypriot Maronite Arabic*

whereas the current local population around the village is comprised of Turkish speakers (many of whom also know Greek, but no longer use it as a language of public life).

Since 1974 the permanent population of the village of Kormakiti has amounted to at most a couple of hundred residents, with the rest of the Maronite community residing primarily in the capital city Nicosia. The Maronite community has occupied a special place in Cypriot society, as for three decades they alone had the ability to freely cross the UN-monitored “Green Line” (buffer zone) dividing the island. Thus connections with the village have been maintained throughout this period, and weekend visits are common. Since 2003 the line has been crossable for all Cypriots.

1.2 Current situation of Cypriot Maronite Arabic

The Cypriot Maronite community currently numbers roughly 5,000 individuals. However, only approximately one thousand are CyA speakers (estimates range from 900 to perhaps 1300; [Council of Europe 2017](#)).

All CyA speakers are bilingual in Cypriot Greek, with Greek as their dominant language, and currently living in a heavily Greek-dominant urban area. There are currently no fluent native speakers under the age of thirty. Due to these factors, the CyA language was designated as severely endangered by UNESCO in 2002.

However, the accession of the Republic of Cyprus to the European Union in 2004 has led to an influx of both institutional and financial support for CyA. In its 2004 initial report on its implementation of the European Charter for Regional or Minority Languages (ECRML), which it ratified in 2002, the Republic of Cyprus declared Armenian as such a language in Cyprus. Although CyA was explicitly excluded as being “only” a dialect and therefore in no need of protection, this formulation was not accepted by ECRML, and CyA was thenceforth officially recognized as a minority language of Cyprus as well. Since 2008 Maronites have been officially recognized as a separate community within Cyprus, and are no longer required to identify themselves as Greek Cypriots (or Turkish Cypriots) on government documents.

The change in designation of the Cypriot Maronites as a linguistic as well as religious minority community led to associated changes in the linguistic rights legally accorded to them. After decades of waiting, one state school in Nicosia is now designated as Maronite and offers optional after-school classes in CyA for its approximately 100 Maronite students, the majority of whom have now joined the classes. Adults may also study CyA now at the new community center. Funding was also made available for a one-to-two week summer language

Mary Ann Walter

immersion camp for Maronite youth in Kormakiti village, attendance at which has risen to approximately 100. For the first time, training seminars for teachers have also been organized, concomitantly with codification efforts towards a written version of CyA. Sporadic writing in CyA has been carried out using the Greek alphabet. (See the community websites in the Further Reading section at the end of this chapter).

Outside the government, there is also an NGO *Hki fi Sanna* ('speak in our language') with the goal of promoting CyA use. Usage remains community- and home-based, as standard Greek (and English) is the language of written and broadcast media. The Cyprus Center of the Peace Research Institute Oslo (PRIO) has undertaken a project entitled *The Protection and Revival of Cypriot Maronite Arabic*. The scope of the project included a variety of community activities, as well as meetings with Sami (Norway) community members for sharing revitalization strategies, described in the resulting publication (PRIO 2009). Finally, a project at the University of Cyprus titled *The Creation of an Archive of Oral Tradition for Cypriot Maronite Arabic* is currently underway under the supervision of Dr. Marilena Karyolemou, though with no web presence or published deliverables to date. There is thus some reason for optimism regarding the future of CyA.

2 Contact languages

CyA has undergone intensive language contact with Cypriot Greek for the entirety of its presence in Cyprus, which may extend to a millennium (see §1.1). This contact has intensified since the removal of the population from the traditionally Maronite and CyA-speaking village of Kormakiti to the capital city, Nicosia.

This move has also resulted in a concomitantly larger social role for Standard Greek. Cyprus is a diglossic society in which Cypriot Greek coexists with Standard Greek, the language of education and formal domains.² In moving to Nicosia, the children of the community also began attending schools with Greek Cypriot children, rather than their own village schools. Only in the last few years has a primary school been designated specifically for Maronite children. Most of them still attend other schools, and the Maronite school is in any case also (Standard-)Greek-medium and follows the same national curriculum (with the addition of optional after-school weekly CyA language classes).

²Some in fact refer to triglossia, encompassing Standard Greek, koineized Cypriot Greek, and various other local varieties, with the island-wide koine taking a mesolectal position (Arvaniti 2010).

3 Cypriot Maronite Arabic

Therefore, the influence of Greek has increased radically through contact with Greek classmates and neighbors, as well as intermarriages with Greek Cypriots and Maronites from other, non-CyA speaking villages. Such situations are common due to the small size of the Maronite community, and typically CyA is not used in these households.

In comparison, contact with Turkish has been limited. Although remaining residents of Kormakiti are now surrounded by Turkish speakers, the village remains quite set apart socially, to the extent that all water supplies are trucked in rather than plumbing systems being shared. In Borg's (1985) texts, speakers do mention using Turkish with some speakers employed as farm workers, however. Contact with Turkish speakers in Nicosia is, of course, rare.

Cyprus is "double-diglossic": the same situation as with Cypriot and Standard Greek holds also with respect to Cypriot and Standard Turkish. To the extent that contact with Turkish does occur, it is with Cypriot Turkish rather than Standard Turkish, unlike Greek, where both varieties are prominent in the lives of CyA speakers.

There is next to no contact with other varieties of Arabic. The Maronite clergy in Cyprus often come from Lebanon, and some intermarriage occurred in the more distant past between the Cypriot and Lebanese Maronite communities, but this no longer occurs. Roth (2004) refers to the "double minoritization" of CyA speakers with respect to both the Cypriot context and the wider Arabophone context – in both, their speech variety is considered deviant and unintelligible.

While early research on CyA identifies it as a Levantine variety of Arabic (Tsiapera 1969), Borg (1985; 2004) argues strongly for an Anatolian origin with significant Aramaic substrate influence. Because the Aramaic influence, if any, must have occurred in the pre-Cyprus period, contact with Aramaic will not be considered further here, despite its putative influence. A substantial discussion can be found in Borg (2004).

Another Semitic language, Syriac, is the liturgical language of the Maronite community. However, no instruction is available in Syriac in Cyprus, so its use is limited to rote recitation during (very sparsely attended) church services, at which transliterations and Greek translations are also provided.

English is the third official language of the Republic of Cyprus (along with Greek and Turkish) and is widely spoken. Instruction in English begins in primary school in the national curriculum, and private English-medium schools are also widespread. However, contact with English postdates contact with Greek and Turkish (beginning only after 1878 and intensifying in the twentieth century) and appears to have had no effects on CyA language structures.

Mary Ann Walter

The French school in Nicosia is traditionally a popular choice for Maronite families, so that competence in French has also been common in the community – a shared characteristic with Lebanese Maronite society. However, like English, this appears not to have influenced CyA grammar in any significant way.

Remaining minority languages of Cyprus include Armenian and a variety of Romani locally called *Kurbetça/Gurbetça*. The reports of ECRML specify that there has been no contact requested or arranged between the Armenian and Maronite community institutions, however. The small size of the communities (each less than 1% of the population) no doubt also reduces the chances of contact. As for *Kurbetça*, it is unclear whether or not it is still actually spoken on the island. Members of this community are Turkish-speaking and interact little if at all with the Maronite community.

Finally, the most common immigrant language (after English) is Russian, which occupies an increasingly prominent place in the linguistic landscape of Cyprus. There are now several Russian-medium schools on the island. However, these are primarily located outside the capital, and its recent appearance means that it also has not influenced CyA.

Therefore, the next section will focus on contact effects from Greek on CyA.

3 Contact-induced changes

According to Borg, the doyen of CyA studies, “linguistic acculturation to Greek in [CyA] is fairly extensive...and involves transfer of allophonic rules, function words, and virtually unrestricted borrowing of content words in the context of codeswitching” as well as “a significant degree of calquing on Greek idioms” (2004: 64). This occurs to such an extent that he describes CyA as “Greek in transparent Arabic garb”, although “the degree of hellenization...tends to be concealed...the inflectional pattern of [CyA] having largely resisted significant intrusion of Greek morphological elements” (2004: 65).

In the remainder of this section, we will examine examples of such Greek influence, particularly in the phonological domain. At the same time, the remarkable persistence of CyA language patterns in the face of intensive contact, especially in the morphological domain, will be discussed.

3.1 Phonology

CyA phonology has been heavily restructured in comparison with other varieties of Arabic, resulting in what Roth (2004: 55) calls “total convergence” of the phono-

3 Cypriot Maronite Arabic

logical system with Cypriot Greek. Similarly, Gulle (2016: 47) refers to the “complete adoption of Greek phonology.”

Like other varieties of Arabic in intense contact with non-Semitic languages, CyA has lost the series of so-called emphatic, guttural or pharyngealized consonants. The obstruents have merged with their non-emphatic counterparts, while the pharyngeal fricative *ħ* has merged with the original glottal fricative *h*, which in turn is now pronounced as a velar fricative [x] under the influence of Greek, as in the examples in Table 1.³

Table 1: Reflexes of emphatic and guttural consonants in CyA

CyA	Arabic	Gloss
<i>taraf</i>	<i>ṭaraf</i>	‘end’
<i>txin</i>	<i>ṭaḥīn</i>	‘flour’
<i>pakar</i>	<i>baqar</i>	‘cattle’
<i>axsen</i>	<i>aḥsan</i>	‘better’

The sole survivals among the Arabic consonants without counterparts in Greek are the interdental consonants and the pharyngeal glide /ʕ/ (see example 5b below). It is interesting that the pharyngeal glide, perhaps the most typologically unusual, remains as a sort of iconic survivor of the Arabic phonemic inventory. The retention of this phoneme, alongside the loss of so many others, implies that the radical changes to the consonant inventory of CyA, though clearly linked to Greek influence, cannot be wholly attributed to imposition in the sense of Van Coetsem (1988; 2000) – or at least, is evidence of significant resistance to such imposition. In any case, imposition would presumably be due to late learners of CyA, and it is doubtful that CyA was ever acquired in this way by speakers from outside the community.

As for the vowels, the Arabic vowel length contrast has also been lost, unstressed (formerly) short vowels deleted, and mid vowels have joined the inventory, resulting in a 5-vowel inventory matching that of Greek, as illustrated in Table 2.

This unsurprising result also occurred in other contact varieties such as Maltese and Andalusi Arabic, although may have evolved without the influence of contact, as in some Levantine varieties.

³Examples are taken from Borg’s (2004) glossary except where noted otherwise. CyA forms are given in his orthography. “Arabic” forms are the presumed etymological source forms, typically shared by standard Arabic as well as other varieties.

Mary Ann Walter

Table 2: Illustration of the innovative vowel system of CyA

CyA	Arabic	CyA Gloss
<i>ipn</i>	<i>ibn</i>	‘son’
<i>umm</i>	<i>umm</i>	‘mother’
<i>tarp</i>	<i>turāb</i> ‘dust’	‘road’
<i>klep</i>	<i>kilāb</i>	‘dogs’
<i>xtop</i>	<i>ktub</i>	‘write’
<i>ten</i>	<i>yadayn</i>	‘hands.DU’

Phonotactically speaking, CyA remains more permissive than Cypriot Greek, in that it “allows a wider range of final consonants and is alone [relative to Cypriot Greek] in allowing final clusters” (Newton 1964: 51).

The effect of (Cypriot) Greek has not been limited to the phonemic inventory. CyA also conforms in the realm of alternations. Like Cypriot Greek, CyA has absolute neutralization of voicing in stop consonants, as illustrated in Table 3.

Table 3: Voicing neutralization in CyA stop consonants

CyA	Arabic	Gloss
<i>sipel</i>	<i>sabal</i>	‘stubble’
<i>ʃates</i>	<i>ʃadas</i>	‘lentils’
<i>pakar</i>	<i>baqar</i>	‘cattle’

It also has the same palatalization and spirantization rules (with the latter applying to the first member of consonant clusters), as well as epenthesis of transitional occlusives in clusters (Tsiopera 1969; Borg 1985; Roth 2004), as illustrated in Table 4.

Table 4: Greek-derived phonological processes in CyA

CyA	Arabic	Gloss	Phonological process
<i>kʲilp</i>	<i>kalb</i>	‘dog’	Palatalization
<i>xtuft</i>	<i>katabt</i>	‘I wrote’	Spirantization
<i>pkyut</i>	<i>buyūt</i>	‘houses’	Consonant epenthesis

3 Cypriot Maronite Arabic

As with changes in the phoneme inventory, these additions to the phonological rules of CyA imply considerable L2 pronunciation effects of Cypriot Greek, even though it was presumably typically acquired later in life than CyA, a puzzling apparent contradiction.

3.2 Morphology

According to Newton (1964: 43), “words of Arabic [...] origin retain the full morphological apparatus of Arabic while those of Cypriot-Greek [...] origin appear exactly as they do in the mouths of monolingual speakers of the Greek dialect.” He goes on to state that “the exceptions to the rule that the morphemes of any one word are either exclusively [Cypriot Greek] or exclusively [Arabic] in origin would seem to be few,” and that Greek verbs “are conjugated exactly as they are when they occur in [Greek].” Example sentences that he provides contain multiple code-switches between Arabic and Greek-origin words, as in (1), where Greek words are highlighted in bold.

- (1) CyA (Newton 1964: 49)
 paxsop **na enicaso** xamse **kamares**
 intend.IMPF.1SG SBJV rent.PRS.1SG five room.PL
 ‘I intend to rent five rooms.’

Newton (1964: 50) concludes that neither source “would be in a position to claim an undisputed majority [of words/morphemes].” Gulle (2016) also discusses examples of “loss of systemic integration” morphologically, with respect to noun plurals, meaning that Greek-origin nouns are used with Greek affixal morphology rather than being integrated into the CyA morphological system. The example in (2) illustrates the use of Greek-origin nouns with Greek plural morphology intact (in bold) in a CyA matrix sentence.

- (2) CyA (Borg 1985: 183, 193)
 allik **p-petrokop-i** n-tammet **l-ispiriðk-ya** ta
 DEM.PL DEF-stonecutter-PL PASS-end.PRF.3SG.F DEF-match-PL COMP
 kan-yišelu **fayy-es**
 PROG.PST-light.IMPF.3PL dynamite.hole-PL
 ‘While those stonecutters were igniting sticks of dynamite, the matches got used up.’

On the whole, the picture is of a language somewhat similar to Maltese (see Lucas & Čěplö this volume), in that we have two morphological systems oper-

Mary Ann Walter

ating in parallel, depending on the etymological origin of the root (Romance or Arabic, in the case of Maltese; Greek or Arabic, in the case of CyA). Alternatively, we could say that speech in CyA is replete with code-switching, and the use of such Greek forms says nothing about the system of CyA itself.

The main exception to morphological non-interaction between CyA and Greek is the use of the Greek diminutive suffix *-ui* (feminine *-ua*) with native CyA words, noted by all three of the major authors on CyA (Borg, Tsiapera, and Newton). For example, this suffix is used with Arabic nouns such as *xmara* ‘female donkey’ and *pint* ‘girl’, yielding *xmarua* ‘small donkey’ and *pindua* ‘girl’ (Newton 1964: 43–44). Tsiapera (1964) additionally notes the borrowing of two adjectival suffixes, *-edin* (which makes nouns into adjectives) and nominal masculine singular *-o*.

Relatedly, Gulle (2016) observes that CyA lacks marking for directive and locative, unlike other Arabic varieties but like spoken Greek. Accusative case marking is used in spoken Greek for this purpose, but due to the lack of overt case marking in CyA, such constructions are unmarked entirely.

(3) CyA (Gulle 2016: 44)

- a. k-kafene
DEF-cafe
‘(in) the cafe’
- b. fi-l-lixkali
in-DEF-field
‘in the field’

Occasional use of Arabic *fi* ‘in,’ as in other varieties and example (3b), was attributed by some CyA speakers of Gulle’s acquaintance to the influence of Levantine Arabic. For at least one speaker, the usage of locative/directional *fi* appeared to be influenced by calquing from standard Greek.

However, Borg (2004: 3) notes similar usage in Old Arabic and Hebrew, such that Greek is not necessarily the source of this pattern. Gulle (2016: 47) concludes that “the tense-aspect-modality (TAM) system [of CyA] is surprisingly almost completely intact”, adding only the exception of the use of the Greek modal verb *prepi* in necessitative constructions.

Finally, the occasional borrowing of the Greek plural morpheme is observed. However, this is sporadic, and a quantitative investigation of pluralization based on Borg’s (2004) glossary (Walter 2017) reveals that native non-suffixal plurals are still used for over half of all pluralizable nouns, at percentages even higher

3 Cypriot Maronite Arabic

than those posited for other Arabic varieties. Greek plurals were given for only 8 of the 251 nouns.

Therefore, although the typically-Arabic use of non-concatenative plural morphology is indeed subject to some degree of suffixal regularization (17% of cases) and somewhat more restricted in terms of the variety of plural forms in CyA, the effect of Greek plural forms has been negligible.

Plural formation, perhaps the most distinctive and cross-linguistically idiosyncratic morphological characteristic of CyA, thus appears remarkably robust in the face of contact. This echoes the retention of the pharyngeal glide in the phonological domain.

On the whole, as Borg (1994: 57) states, “the external impact on the native morphological patterns of [CyA] is slight.”

3.3 Syntax

According to Roth (2004: 70), “syntax is a linguistic domain particularly permeable to interference from Greek” (author’s translation). By this she means that function words are doubled with loans from Greek, in particular with relative clause markers and more complex constructions, as well as the use of Greek and Arabic-origin negation markers in combination. The example in (4) demonstrates the CyA use of the native *ma* negation morpheme concurrently with Greek *me...me*. In this case, phonetic similarity may have aided the adoption of *me*.

- (4) a. CyA (Borg 1985: 149;
)ma-pišrap me pira me mpit
 NEG-drink.IMPF.1SG NEG beer NEG wine
 ‘I don’t drink either beer or wine.’
 b. (Cypriot Greek)
 em-pinno me piran me krasin
 PROG-drink.PRS.1SG NEG beer.ACC NEG wine.ACC
 ‘I don’t drink either beer or wine.’

It is unclear, however, whether all or most of this is simply codeswitching and whether it should be termed syntactic rather than lexical influence.

A syntactic change which does not involve codeswitching or lexical borrowing is the development of a predicative copula (lacking in the present tense in most varieties of Arabic) from Arabic pronouns, discussed by both Roth (2004) and Borg (1985), and illustrated in (5).

Mary Ann Walter

- (5) CyA (Borg 1985: 134)
- a. l-iknise e maftux-a
DEF-church 3SG.F open-F
‘The church is open.’
 - b. p-pkyara enne maʕak
DEF-well.PL 3PL deep.PL
‘The wells are deep.’

In example (5a), the copula corresponds to the third-person feminine pronoun ‘she’ (also *e*, < *hiya*). Likewise, the copula *enne* in (5b) corresponds to the third-person plural pronoun ‘they’ (also *enne*; < *hunna*). The development of this copula presumably replicates the obligatory present-tense copula of Greek. See Lucas & Čéplö (this volume) for a similar phenomenon in Maltese.

Finally, both Roth (2004) and Newton (1964) document variable placement of adjectives, according to both Arabic and Greek norms, as illustrated in (6).

- (6) a. CyA (Roth 2004: 72)
m-mor-a li-zʕar
DEF-child-PL DEF-small.PL
‘small children’
- b. Lebanese Arabic (Newton 1964: 48)
l-bēt l-ikbīr
DEF-house DEF-big
‘the big house’
- c. CyA (Newton 1964: 47)
li-kbir payt
DEF-big house
‘the big house’
- d. Cypriot Greek (Newton 1964: 48)
to meálo spítin
DEF big house
‘the big house’

However, Borg (2004) notes that so-called “peripheral” varieties of colloquial Arabic have been said to employ freer word order than others, so the variation in noun–adjective ordering may be an independent internal development (or alternatively, perhaps peripheral varieties are by nature more subject to contact, which leads to this pattern of variation).

3 *Cypriot Maronite Arabic*

In summary, syntax, like morphology, shows relatively little influence of language contact, especially in contrast to the phonological system. As word order is already relatively flexible in both CyA and Cypriot Greek (e.g. with respect to subject–verb ordering; [Newton 1964](#): 48–49), this is perhaps to be expected.

3.4 **Lexicon**

According to [Newton \(1964\)](#), of the 630 common lexical items which he elicited, 38% were Greek in origin. However, he goes on to say that the percentage is lower in running speech, in which typically the most common (and therefore native Arabic origin) vocabulary was used. He raises the possibility ([1964](#): 51) that CyA consists of “Arabic plus a large number of Cypriot [Greek] phrases thrown in whenever [a speaker’s] Arabic fails him or the fancy takes him.” [Tsi-paera \(1964: 124\)](#) concurs, stating that “any speaker of [CyA] has a minimum of about thirty per cent of Greek lexical items in his speech which are not assimilated into the phonological and morphological system of his native language.” She identifies the semantic fields of government and politics, numerical systems including weights and measures, and adverbial particles as particularly dominated by words of Greek origin.

This percentage contrasts with the relatively small number of Greek-origin items appearing in [Borg’s \(2004\)](#) glossary. However, the difference in elicitation contexts must be kept in mind – Newton’s work occurring in the Cypriot context and himself being competent in Greek, versus Borg’s work occurring partly overseas and himself an Arabist rather than a scholar of Greek.

[Roth \(2004\)](#) refers to the drastic reduction of the lexicon, and estimates that it includes at most 1300 items. Borg’s ([2004](#)) glossary contains roughly 2000 entries (corresponding to 720 lexical consonantal roots), which he considers to be a “substantial portion” (though not all) of the “depleted” Arabic-origin CyA lexicon.

[Gulle \(2016: 45\)](#) notes suppletion in the paradigm of the verb ‘to come’, with imperative forms borrowed from Greek. The consonantal root of the verb ‘to come’, in CyA as elsewhere in Arabic, is $\sqrt{\text{žy}}$, as seen in the form *ža* ‘he came’. However, CyA imperative forms of this verb (*ela*, *eli*, *elu*, in masculine singular, feminine singular, and plural forms respectively) are clearly based on Greek *ela*, *elate* (singular and plural, respectively). This particular case seems to reflect a pan-Balkan spread of this item, as *ela/elate* are also used in Bulgarian (personal knowledge).

In summary, universal bilingualism and Greek-dominance among CyA speakers results in widespread use of codeswitched Greek vocabulary and associated

Mary Ann Walter

morphology, with marginal lexical suppletion. However, there is very little loan material integrated into the CyA grammatical system.

As a final note, Hadjidemetriou's (2009) doctoral dissertation examines language contact between CyA and Cypriot Greek (as well as Armenian and Cypriot Greek), in the opposite direction, to identify any effects of CyA on Cypriot Greek. Unsurprisingly, however, given the current dominance of Cypriot Greek for these speakers, no such effects were found, in any of the above domains.

4 Conclusion

CyA appears to present a counterexample to Van Coetsem's notion of the stability gradient, which claims that phonology (and syntax) are more stable than other domains (the lexicon). It is clear that for CyA, phonology has been the least stable domain. The observed phonological convergence to Greek is of the type that suggests pervasive effects of L2 pronunciation (except for the retention of the pharyngeal glide). Yet it is difficult to imagine any sociolinguistic scenario in which CyA was taken up in any significant numbers by Greek speakers from outside the community, and the typical acquisition scenario (when CyA was still acquired by children) has been use of CyA as a home language, and Greek as a school language, thereby generating sequential (though eventually probably Greek-dominant) bilinguals. The historical record is unfortunately lacking any relevant information that could shed light on the situation.

The most urgent issue for future research on CyA is undoubtedly the need for additional documentation efforts. In particular, naturalistic texts and audio recordings are a desideratum. It is to be hoped that the documentation and revitalization efforts currently underway will remedy this situation.

Further reading

Tsiapera's (1969) work is the only one so far to consider CyA in its totality as a spoken language, although not at great length, and it drew subsequent criticism of the author's lack of background knowledge of the Arabic language. This monograph does, however, have the additional advantage of a publication date very close in time to the radical change in the sociolinguistic circumstances of CyA speakers due to ethnic tensions in the island, culminating in their near-unanimous relocation from traditionally Maronite villages to the capital city Nicosia.

3 Cypriot Maronite Arabic

Borg’s (1985) foundational work on morphophonology is still the most extensive resource on CyA grammar. He takes a historical perspective on changes from earlier Arabic to contemporary CyA, both contact-driven and otherwise, and also includes substantial textual material in CyA at the end. These texts are currently the only published ones available.

The follow-up volume by Borg (2004) includes a substantial introductory essay situating CyA within the range of Arabic dialects and elucidating the influences of the main contact language, Cypriot Greek. The lexical entries are enriched by comparisons with dialectal forms from other varieties of Arabic, as well as Greek, Aramaic, and other contact languages where relevant.

The most up-to-date and reliable information regarding CyA and its speakers, including documentation, preservation and revival efforts, may be found in the Council of Europe (2017) report.

The following two community websites contain information on CyA institutions and activities in both Greek and English, including contact information, historical background, archived copies of the monthly (Greek language) community newsletter, and so on.

<http://www.maronitesofcyprus.com> (in both Greek and English)
<http://kormakitis.net/portal/> (in Greek)

Acknowledgements

My sincere thanks to Christopher Lucas and Stefano Manfredi for organizing the workshop on Arabic and language contact at the 23rd International Conference on Historical Linguistics, and their editing work on this volume. I would also like to thank the audience at ICHL for their insightful comments, as well as those of the members of the Processing and Acquisition of Language Lab at Cambridge University (CAMPAL). Remaining errors and infelicities are of course my own.

Abbreviations

ACC	accusative	DEF	definite
CAMPAL	Cambridge Processing and Acquisition of Language Lab	DEM	demonstrative
CE	Common Era	DU	dual
CyA	Cypriot Maronite Arabic	ECRML	European Charter for Regional or Minority Languages
COMP	complementizer		

Mary Ann Walter

F	feminine	PST	past tense
IMPF	imperfect	SBJV	subjunctive
PASS	passive	SG	singular
PRF	perfect	TAM	tense-aspect-modality
PL	plural	UN	United Nations
PRIO	Peace Research Institute Oslo	UNESCO	United Nations Educational, Scientific and Cultural Organization
PROG	progressive		
PRS	present		

References

- Arvaniti, Amalia. 2010. Linguistic practices in Cyprus and the emergence of Cypriot Standard Greek. *Mediterranean Language Review* 17. 15–45.
- Baider, Fabienne & Marilena Kariolemou. 2015. Linguistic unheimlichkeit: The Armenian and arab communities of Cyprus. In Ulrike Jessner-Schmid & Claire Kramsch (eds.), *The multilingual challenge: Cross-disciplinary perspectives*, 185–209. Berlin: De Gruyter Mouton.
- Borg, Alexander. 1985. *Cypriot Arabic: a historical and comparative investigation into the phonology and morphology of the arabic vernacular spoken by the Maronites of Kormakiti village in the Kyrenia district of north-west Cyprus*. Stuttgart: Deutsche Morgenländische Gesellschaft.
- Borg, Alexander. 1994. Some evolutionary parallels and divergences in cypriot arabic and maltese. *Mediterranean Language Review* 8. 41–67.
- Borg, Alexander. 2004. *A comparative glossary of Cypriot Maronite Arabic (Arabic–English): With an introductory essay*. Leiden: Brill.
- Council of Europe. 2017. *Fifth periodical report presented to the Secretary General of the Council of Europe in accordance with Article 15 of the European Charter for Regional or Minority Languages*. Strasbourg: Council of Europe.
- Gulle, Ozan. 2016. Kormakiti Arabic: A study of language decay and language death. In Vera Ferreira & Peter Bouda (eds.), *Language documentation and conservation in Europe*, 38–50. Honolulu: University of Hawai'i Press.
- Hadjidemetriou, Chryso. 2009. *The consequences of language contact: Armenian and Maronite Arabic in contact with Cypriot Greek*. Colchester: University of Essex doctoral dissertation.
- Newton, Brian. 1964. An Arabic–Greek dialect. *Word* 20(sup2). 43–52.
- PRIO. 2009. *The protection and revival of Cypriot Maronite Arabic*. <https://www.prio.org/Publications/Publication/?x=7371>.

3 Cypriot Maronite Arabic

- Roth, Arlette. 2004. Le parler arabe maronite de chypre: Observations à propos d'un contact linguistique pluriséculaire. *International Journal of the Sociology of Language* 168. 55–76.
- Tsiapera, Maria. 1964. Greek borrowings in the Arabic dialect of Cyprus. *Journal of the American Oriental Society* 84(2). 124–126.
- Tsiapera, Maria. 1969. *A descriptive analysis of Cypriot Maronite Arabic*. The Hague: Mouton.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Walter, Mary-Ann. 2017. A quantitative investigation of noun pluralization in cypriot maronite and maltese Arabic. Paper presented at the 23rd International Conference on Historical Linguistics, San Antonio, Texas, July 31 - August 4.

Chapter 4

Maghrebi Arabic

Adam Benkato

University of California, Berkeley

This chapter gives an overview of contact-induced changes in the Maghrebi dialect group in North Africa. It includes both a general summary of relevant research on the topic and a selection of case studies which exemplify contact-induced changes in the areas of phonology, morphology, syntax, and lexicon.

1 The Maghrebi Arabic varieties

In Arabic dialectology, MAGHREBI is generally considered to be one of the main dialect groups of Arabic, denoting the dialects spoken in a region stretching from the Nile delta to Africa’s Atlantic coast – in other words, the dialects of Mauritania, Morocco, Algeria, Tunisia, Libya, parts of western Egypt, and Malta. The main isogloss distinguishing Maghrebi dialects from non-Maghrebi dialects is the first person of the imperfect, as shown in Table 1 (cf. Lucas & Čépľo this volume).¹

Table 1: First person imperfect ‘write’ in Maghrebi and non-Maghrebi Arabic

	Non-Maghrebi		Maghrebi	
	Classical Arabic	Baghdad Arabic	Casablanca Arabic	Maltese
Singular	<i>aktub</i>	<i>aktib</i>	<i>nəktəb</i>	<i>nikteb</i>
Plural	<i>naktub</i>	<i>niktib</i>	<i>nkətbu</i>	<i>niktbu</i>

¹More about the exact distribution of this isogloss can be found in Behnstedt (2016).

Adam Benkato

This Maghrebi group of dialects is in turn traditionally held to consist of two subtypes: those spoken by sedentary populations in the old urban centers of North Africa, and those spoken by nomadic populations. The former of these, usually referred to as “pre-Hilali” (better: “first-layer”) would have originated with the earliest Arab communities established across North Africa (~7th–8th centuries CE) up to the Iberian peninsula. The latter of these, usually referred to as “Hilali” (better: “second-layer”), is held to have originated with the westward migration of a large group of Bedouin tribes (~11th century CE) out of the Arabian peninsula and into North Africa via Egypt. Their distribution is roughly as follows.² First-layer dialects exist in cities such as Tunis, Kairouan, Mahdia, Sousse, Sfax (Tunisia), Jijel, Algiers, Cherchell, Tlemcen (Algeria), Tangier, Tetuan, southern Rif villages, Rabat, Fez, Taza, so-called “northern” dialects (Morocco), Maltese, and formerly Andalusi and Sicilian dialects; most Judeo-Arabic dialects formerly spoken in parts of North Africa are also part of this group. Second-layer dialects are spoken by populations of nearly all other regions, from western Egypt, through all urban and rural parts of Libya, to the remaining urban and rural parts of Algeria and Morocco. Though some differences between these two subtypes are clear (such as [q, ʔ, k] vs. [g] for *q), there have probably been varying levels of interdialectal mixture and contact since the eleventh century CE. In many cases, first-layer varieties of urban centers have been influenced by neighboring second-layer ones, leading to new dialects formed on the basis of inter-dialectal contact. It is important to note that North Africa is becoming increasingly urbanized and so not only is the traditional sedentary/nomadic distinction anachronistic (if it was ever completely true), but also that intensifying dialect contact accompanying urbanization means that new ways of thinking about Maghrebi dialects are necessary. It is also possible to speak of the recent but ongoing koinéization of multiple local varieties into supralocal or even roughly national varieties—thus one can speak, in a general way, of “Libyan Arabic” or “Moroccan Arabic”. This chapter will not deal with contact between mutually intelligible varieties of a language although this is equally important for the understanding of both the history and present of Maghrebi dialects.³

²More will not be said about the subgroups of Maghrebi dialects that have been proposed. For more details about the features and distribution of Maghrebi dialects see [Pereira \(2011\)](#); for more detail on the complex distribution of varieties in Morocco see [Heath \(2002\)](#).

³The emergence of new Maghrebi varieties resulting from migration and mixture is discussed in [Pereira \(2007\)](#) and [Gibson \(2002\)](#), for example. The oft-cited distinction between urban and nomadic dialects is also problematized by the existence of the so-called rural or village dialects (though this is also a problematic ecolinguistic term), on which see [Mion \(2015\)](#). Dialect contact outside of the Maghreb is discussed by Cotter (this volume).

2 Languages in contact

Contact between Arabic and other languages in North Africa began in the late seventh century CE, when Arab armies began to spread westward through North Africa, reaching the Iberian peninsula by the early eighth century CE and founding or occupying settlements along the way. Their dialects would have come into contact with the languages spoken in coastal regions at that time, including varieties of Berber and Late Latin, and possibly even late forms of Punic and Greek. The numbers of Arabic-speakers moving into North Africa at the time of initial conquests was likely to have been quite small.⁴ By the time of the migration of Bedouin groups beginning in the eleventh century, it is doubtful that languages other than Berber and Arabic survived in the Maghreb. The Arabization of coastal hinterlands and the Sahara increased in pace after the eleventh century. Berber varieties continue to be spoken natively by millions in Morocco and Algeria, and by smaller communities in Libya, Tunisia, Mauritania, and Egypt. Any changes in an Arabic variety due to Berber are almost certainly the result of Berber speakers adopting Arabic rather than Arabic speakers adopting Berber – the sociolinguistic situation in North Africa is such that L1 Arabic speakers rarely acquire Berber.

Beginning in the sixteenth century, most of North Africa came under the control of the Ottoman Empire and thus into contact with varieties of Turkish, although the effect of Turkish is essentially limited to cultural borrowings (see below, 3.4). The sociolinguistic conditions in which Turkish was spoken in North Africa are poorly understood.

The advent of colonialism imposed different European languages on the region, most prominently French (in Mauritania, Morocco, Algeria, and Tunisia), Italian (in Libya), and Spanish (in Morocco). Romance words in dialects outside of Morocco may also derive from forms of Spanish (via Andalusí refugees to North Africa in the 16th–17th centuries) or from the Mediterranean *Lingua Franca*.⁵

The effects on Maghrebi Arabic of contact with Chadic (Hausa) or Nilo-Saharan (Songhay, Tebu) languages is largely unstudied since in most cases data from the relevant Arabic varieties is lacking. Yet some borrowings from these languages can be found in Arabic and Berber varieties throughout the region (Souag 2013).⁶ Lastly, Hebrew loans are present in most Jewish Arabic dialects of North Africa (Yoda 2013), though unfortunately these dialects hardly exist anymore.

⁴See Heath (this volume) for discussion of Late Latin influence in Moroccan Arabic dialects.

⁵On the *Lingua Franca* see Nolan (this volume).

⁶See also Souag (2016) for an overview of contact in the Sahara region not limited to Arabic.

Adam Benkato

To restate these facts in Van Coetsem's (1988; 2000) terms, there are two major contact situations at work in Maghrebi Arabic in general, though the specifics will of course differ from variety to variety. The first is change in Arabic driven by source-language (Berber) dominant speakers; this transfer type is IMPOSITION. The second is change in Arabic driven by recipient-language (Arabic) dominant speakers where the source language is a European colonial language; this transfer type is called BORROWING.⁷ So far, "dominance" describes linguistic dominance, that is, the fact that a speaker is more proficient in one of the languages involved in the contact situation. However, social dominance, referring to the social and political status of a language (Van Coetsem 1988: 13), is also important, especially in North Africa.

3 Contact-induced changes in Maghrebi dialects

3.1 Phonology

Changes in Maghrebi Arabic phonology due to contact with Berber are difficult to prove. There are several cases, for example, where historical changes in Arabic phonology may be argued to be the result of contact with Berber *or* the result of internal developments. These include the change of *ǧ to /ʒ/ in many varieties, or the emergence of phonemic /z/ (Souag 2016). Another example, the pronunciation /t/ in some first-layer varieties where most Arabic varieties have /ð/ has also been explained as a result of Berber influence, or as unclear directionality (Kossmann 2013b: 187), while Al-Jallad (2015) argues that it is actually an archaism within Arabic.

The merger in Arabic of the vowels *a and *i (and even *u) to a single phoneme /ə/ in some, especially first-layer, varieties, is often attributed to Berber influence, as many Berber varieties have only a single short vowel phoneme /ə/. However Kossmann (2013b: 171–174) points out that Berber also merged older *ā and *ā to a single phoneme /ə/ and that it cannot be proven that the reduction happened in Berber before it happened in Arabic. Hence, again the directionality of influence is difficult to show.

Related to this development is also that many Maghrebi varieties disallow vowels in light syllables (often described as the deletion of short vowels in open syllables), such that **katab* 'he wrote' > Tripoli *ktəb* or **kitāb* 'book' > Algerian *ktāb*.⁸

⁷Another good illustration of the two transfer types in the Van Coetsemian framework can be found in Winford (2005: 378–381).

⁸Since the short vowels merge to schwa in many Moroccan and Algerian varieties, vowel

Meanwhile, second-layer varieties often do allow vowels in light syllables (e.g. Benghazi *kitab* ‘he wrote’, Douz *mⁱšē* ‘he went’). While proto-Berber and some modern varieties allow vowels in light syllables, most Berber varieties of Algeria and Morocco do not. This is another example of a similar development wherein the directionality of influence is unclear (see Souag 2017: 62–65 for further discussion).

In the Arabic variety of Ghomara, northwest Morocco, *d and *t are spirantized to /ð/ and /θ/ initially (*d only), postvocally and finally (Naciri-Azzouz 2016): e.g. *māθθ* ‘she died’ (*mātat), *warθ* ‘inheritance’ (although etymologically *warθ, dialects of the wider Jbala region of Morocco have no interdentals so *wart), *ḏāba* ‘now’ (*dāba), *ḥaḏma* ‘work’ (*ḥidma), *wāḥaḏ* ‘one’ (*wāhid). Naciri-Azzouz points out that the distribution of spirantization is the same as in Ghomara Berber, a variety spoken by groups in the same region.⁹

New phonemes have been borrowed into Maghrebi varieties through contact with European languages: for example, /p/ and nasalized vowels in more recent French loans in Tunisian Arabic, or /v, č, ğ/ in Italian loans in Libyan Arabic (*grīḡū* ‘gray’ < *grigio*).

3.2 Morphology

In the realm of morphology, changes in Arabic varieties due to contact vary depending on whether the relationship between Arabic and the contact language is substratal, adstratal, or superstratal.

Morphological influence from Berber on Arabic varieties of the northern Maghreb is not overly common.¹⁰ In some places where Berber/Arabic bilingualism is or was more common, contact has led to the borrowing of Berber nouns into Arabic together with their morphology, a phenomenon known as “parallel system borrowing”.¹¹ In Ḥassāniyya, for example, many nouns have been transferred together with their gender and number marking.¹² In the dialect of Jijel,

length is no longer contrastive and it is common to transcribe e.g. *ktab* rather than *ktāb*.

⁹The Berber variety of Ghomara exhibits an extreme amount of influence from dialectal Arabic, see Mourigh (2015). Kossmann (2013b: 431) writes that given the existence of parallel morphological systems for virtually all grammatical categories (nominal, adjectival, pronominal and verbal morphology) and a high loanword count (more than 30% of basic lexicon is Arabic) it would be possible to call Ghomara Berber a mixed language.

¹⁰Documentation of the varieties where such influence would be more expected, such as Arabic-speaking towns in the otherwise Berber-speaking Nafusa Mountains in Libya, is lacking.

¹¹For a closer look at parallel system borrowing in the context of Arabic and Berber contact, see Kossmann (2010), mostly discussing the borrowing of Arabic paradigms into Berber.

¹²See Taine-Cheikh (this volume).

Adam Benkato

Berber singular nouns are transferred together with their prefixes (*āwtūl* ‘hare’, cf. Kabyle *āwtūl*); plurals are then formed in a way which resembles Berber but is not identical (Jijel *āsraḥ*, *āsraḥn* ‘bushe(s)’, cf. Kabyle Berber *āsraḥ*, *īsraḥn*); moreover, the prefix *ā-* is also used with nouns of Arabic origin (*āfhad* ‘thigh’, Arabic **faḥad*) (Marçais 1956: 302–318).

In Algeria and Morocco the circumfix *tā-...-t*, which occurs on feminine nouns in Berber, can derive abstract nouns (e.g. Jijel *tākabūrt* ‘boasting’, *tawḥḥūnt* ‘having labor pains’) and in Moroccan Arabic *tā-...-t* is the regular way of forming nouns of professions and traits (e.g. *tānāžžārt* ‘carpentry’) (Kossmann 2013a).

The verbal morphology of Arabic dialects is much less affected by Berber, though Ḥassāniyya again provides an interesting example. It has a causative prefix *sā-* used with both inherited Arabic verbs and borrowed Berber verbs, and most likely to be borrowed from Berber causative forms in *s-/š-* (Taine-Cheikh 2008).

Turkish influence on morphology is restricted to the suffix *-ḡi/-ḡi* (< *-ci*) used to indicate professions and borrowed widely into Arabic dialects in general. In Tunisia, its use has been extended to derive adjectives of quality from nouns (*sukkārḡi* ‘drunkard’) and has also even been added to borrowed French nouns (*bankāḡi* ‘banker’ < French *banque*). As Manfredi (2018: 410) points out, the productivity of this borrowed derivational morpheme is one example of how recipient-language agentivity can introduce morphological innovations based on borrowing.

French (and other Romance) verbs are also routinely borrowed into Maghrebi varieties. Talmoudi (1986) discusses their integration into different forms of the verbal system of Tunisian Arabic, e.g. *mannak* ‘to be absent’ < French *manquer* or (*t*)*rānā* ‘to train’ < French *trainer*.

3.3 Syntax

Syntax is often the least documented aspect of the grammar of Maghrebi Arabic varieties and research on contact-induced changes in syntax is still in its infancy. Much attention has been devoted recently to explaining the rise of bipartite negation in Arabic and Berber; in varieties of both languages the word for ‘thing’ (Arabic *šayʔ*, Berber **kāra*) has been grammaticalized postverbally in a marker of negation:

- (1) Arabic (Benghazi)

mā-šift-hā-š

NEG-see.PRF.1SG-3SG.F-NEG

‘I didn’t see her’.

(2) Berber (Tarifit)

wā t-ẓriy

ša

NEG 3SG.F-see.PRF.1SG NEG

‘I didn’t see her’.

Although some accounts give no attention to Berber, while others attribute the Arabic development solely to Berber, the development in both languages in the same contexts is probably not a coincidence; though there is no current consensus on the direction of transfer – see Lucas (this volume) for discussion.¹³ However, it must be noted that not all Berber varieties have double negation (e.g. Tashelhiyt *ur nniy ak* ‘I didn’t tell you’ where the only negator is *ur*).

In another area, recent work on the variety of Tunis has yielded interesting conclusions: while possessives with French nouns are overwhelmingly analytic (*l-prononciation mtēf-ha* ‘her pronunciation’) and those with Arabic nouns are almost as overwhelmingly synthetic (*nutq-u* ‘his pronunciation’), the frequent occurrence of French loan nouns may be triggering an increase in the overall frequency of analytical possessives over syntactic ones, including those with Arabic nouns (Sayahi 2015).

The remainder of this section will discuss one particularly interesting case: the first-layer dialect of Jijel, a city in eastern Algeria. At the time of its description (Marçais 1956), it showed little influence from second-layer varieties, but displayed wide-ranging influence from Berber in multiple domains. In a recent article, Kossmann (2014) has demonstrated how a Berber marker of non-verbal predication was adopted into the Arabic dialect of Jijel as a focus marker. Here I will briefly summarize Kossmann’s arguments with a few examples. In the Jijel dialect, as described by Marçais and reanalyzed by Kossmann, a morpheme *d* occurs in the following syntactic contexts (examples (3)–(7) are all from Kossmann 2014: 129–131, who retranscribes from Marçais’ texts): before non-verbal predicates (3), in clefts with a noun/pronoun in the cleft (4), in secondary pred-

¹³See Lucas (2007; 2010; 2018) and Souag (2018) for further discussion of the grammaticalization of ‘thing’ for indefinite quantification and polar question marking in Arabic and Berber. Kossmann (2013b: 324–334) surveys the situation in the Berber languages. See Lafkioui (2013) for an overview of negation in especially Moroccan Arabic, as well as discussion of a variety of Moroccan Arabic which features the discontinuous morpheme *mā- ... -bū*, where the latter part has been borrowed from Tarifit.

Adam Benkato

ication with a specific noun (5), as a marker of subject (or object) focus (6), and in left-moved focalizations (7).

- (3) l-lila d-əl-ʕid
DEF-night D-DEF-feast
'Tonight is the feast.'
- (4) d-hum əddə šraw-əh qbəl-ma nəzɖad
D-3PL.M REL buy.PRF.3PL.M-3SG.M before-COMP be.born.IMPF.1SG
'It is them who bought it before I was born.'
- (5) ɾa-na nqəttʃu-č d ət-ɾaf
PRST-1PL cut.IMPF.1PL-SG D DEF-pieces
'We will cut you (into) pieces.'
- (6) tkəʃʃɾət d l-idura
break.PRF.3SG.F D DEF-bowl
'The bowl has broken.'
- (7) qalu d əɾ-ɾbiʃ dəhlət
say.PRF.3PL D DEF-spring enter.PRF.3SG.F
'They say spring has come.'

Although previous analyses attempted to explain *d* within Arabic, Kossmann notes that an Arabic-internal derivation of *d* is impossible. However, Kabyle, the Berber language neighboring the Jijel area has an element *d* (realized [ð] due to spirantization in Kabyle) which is used in (pro)nominal predicates (8), cleft constructions (9), and secondary predication when non-verbal (10). Examples (8)–(10) are all Kabyle Berber, taken from Kossmann (2014: 135–136). This element *d* is attested in Berber more widely, too, and is likely reconstructible to older stages of the language.

- (8) d-yəlli-m
D-daughter-2SG.F
'Is it your daughter?'
- (9) d-ay-ən i d-tənna abrid amənzɖ
D-this-DEICT REL hither-say.3SG.F road first
'This is what she said the first time.'

- (10) ad nəǧʔəl iman nn-əy d-inəbgiwən n ɾəbbi
 MOD make.1PL self GEN-1PL D-guests GEN lord
 ‘We shall pretend to be beggars (lit. guests of God).’

Thus Berber *d* is the best candidate for the origin of Jijel Arabic *d*, though its usage in (Kabyle) Berber (where it is a primarily marker of syntactic organization) differs from that of Jijel Arabic (where it is mainly a marker of information structure). In a simplified scenario with a Berber variety as source language and Jijel Arabic as recipient, *d* would likely have been imposed into Jijel Arabic with its exact Berber functions. As Kossmann notes, though, speech communities are full of variation and language contact is a “negotiation between the frequency of non-native speech and the prestige of the native way of speaking” (Kossmann 2014: 138). Kossmann thus proposes a scenario in which larger groups of Berber speakers switched to a variety of Jijel Arabic and began imposing their own *d*; the native Jijel Arabic speakers, fewer in number, began adopting *d* but understood it differently and interpreted it as a focus marker, introducing it into new contexts; eventually the variety of Jijel Arabic with *d* in all these functions became nativized. Per Kossmann (2014: 138–139), two processes would have taken place: the transfer of a source-language feature by speakers dominant in the source language (Berber), followed by the borrowing of this feature by speakers dominant in the recipient language (Arabic), and its eventual regularization in that variety. Jijel Arabic is an excellent example of what may happen when large numbers of Berber speakers switch to Arabic.

3.4 Lexicon

Much work on contact and Maghrebi Arabic has focused on loanwords, the most salient effects of borrowing, with secondary attention to their phonological or morphological adaptation. The concept of social dominance has particular relevance for borrowing: in the North African context, the colonial languages, especially French, have high social status for both Arabic and Berber native speakers. One also must modify the idea of linguistic dominance to include those who acquire two languages natively (2L1 speakers; see Lucas 2015: 525), definitely the case for certain speakers of Berber and Arabic in North Africa.

Unsurprisingly, we see firstly that the majority of words borrowed into Arabic varieties are nouns, and secondly that the lexical domains into which these borrowings fall are often restricted. Social dominance seems to play a role in the nature of the nouns borrowed.

Adam Benkato

Berber loans are found in most Maghrebi varieties, though their number ranges from only a handful of words in the east to many more in the west (cf. §3.2 above). Almost all Maghrebi varieties borrow the words *ž(i)rāna* ‘frog’ and *fakrūna* ‘turtle’, while in some oases Berber influence in agricultural terminology can be seen. Again, the documentation of the relevant varieties is often insufficient.

Several studies on contact between Maghrebi Arabic varieties and European languages exist. For French in Morocco, [Heath \(1989\)](#) argues that code-switching and borrowing are essentially the same in a bilingual community which has established borrowing routines.¹⁴ For French in Tunisia, [Talmoudi \(1986\)](#) analyzes the phonological and morphological adaptation of French verbs into Arabic. [Sayahi \(2014: 127–151\)](#) gives a broader view of lexical borrowing in diglossic/bilingual communities, focusing on French in Tunisia and Spanish in Morocco. [Vicente \(2005\)](#) studies Arabic-Spanish code-switching in Ceuta, a Spanish enclave in northern Morocco. Italian in Tunisia is studied briefly by [Cifoletti \(1994\)](#). Studies of contact with Turkish are limited to discussion of loanwords: on Morocco see [Procházka \(2012\)](#); on Algeria, see [Ben Cheneb \(1922\)](#), to be read with the review by G.S. Colin ([Colin 1999: 21–30](#)).

The remainder of this section will consider the influence of Turkish and Italian on Libyan Arabic (henceforth LA), a hitherto under-researched topic. Uniquely in the Maghreb region there is at present no superstratum language spoken widely by Arabic speakers in Libya, while there are also fewer Berber speakers than in Algeria or Morocco. As far as documented varieties of LA (Tripoli and Benghazi) go, contact situations are historical and not active.

There seems to be an impression among dialectologists that LA varieties have the largest number of Turkish loans, though there is not a published basis for this. [Procházka \(2005: 191\)](#) suggests that the number of (Ottoman) Turkish loans in a given Arabic dialect is proportional to the length and intensity of Ottoman rule. By this criterion Libya should have quite a few, as the regions now constituting Libya were under control of the Ottoman Empire from 1551 to 1911, but [Procházka](#) estimates that the dialect would show 200 to 500 surviving loans, fewer than in other dialects. Another important factor is likely to be that Libya’s population

¹⁴[Van Coetsem \(1988: 87\)](#) notes that for bilingual speakers who have a balance in linguistic dominance between the two languages, the separation between the two transfer types (borrowing and imposition) will be weaker. Hence, either of the two dominant languages can serve as the recipient language in codeswitching behavior. [Winford \(2005, esp. 394–396\)](#), expanding on Van Coetsem’s framework, points out that code-switching is inherently linked to the borrowing transfer type. In the Maghreb, this scenario is possible for Berber–Arabic bilinguals as well as for some French–Arabic bilinguals. See [Ziamari \(2005\)](#) for an insightful and more recent analysis of Moroccan Arabic in contact with French using a “matrix language frame” analysis.

4 Maghrebi Arabic

was very small during the period of Ottoman rule so that the long-term presence of even a few thousand Turkish speakers could have had a significant effect. However, I cannot yet offer a statistical analysis of Turkish words in Libyan Arabic.¹⁵ It is clear so far, though, that the effects of Turkish on LA can mainly be seen in the lexicon and, in my data, almost entirely in nouns. In terms of their semantic domains, Procházka (2005: 192) points out that the majority of Turkish loans in Arabic dialects in general fall into three categories, roughly described as: private life; law, government, social classes; and army, war. By far the majority of surviving loans would belong to the first of these classes (such as *šišma* ‘tap’ < *çeşme*, *dizdān* ‘wallet’ < *cüzdan*), or second (such as *fayramān* ‘order’ < *ferman*, *haḫḫa* ‘week’ < *haḫte*), while I suspect that words from the third class are increasingly rarer. Outside of these, only a few words other than nouns seem to be present, such as *duḡri* ‘straight ahead’ and *balki* ‘maybe’. The length of time since Turkish was last actively spoken in Libya no doubt means that the number of Turkish loans actively used by speakers has been decreasing.

LA is unique among Maghrebi varieties in having had Italian as the main European contact language. Italian had a presence in what is now Libya from the 1800s, but this was mainly limited to the Tripolitanian Jewish community and wealthy merchant families. The Italian colonization of Libya officially began in 1911; though the majority of the region was not brought under Italian control until the early 1930s, large numbers of Italian colonists had begun to settle in Libya in the 1920s. From that period until 1970, when the remaining Italian citizens were expelled from the country, Italians made up 15% or more of the population and the language was in widespread use. From the 1970s on, Italian was scarcely used in Libya, and the teaching of foreign languages was banned in 1984, not to return again until 2005.¹⁶ Many of the postwar generation spoke (and still speak) Italian, though they rarely use it anymore, but few Libyans of younger generations do. The 1920s to the 1970s can thus be regarded as the main period of contact between LA and Italian.¹⁷ However, the concentration of Italians differed from region to region and thus may have influenced local varieties

¹⁵The only study dedicated to Turkish loans in LA is Türkmen (1988), who lists 90 words. However, the basis for his wordlist seems unclear and several items are either spurious or incorrect (e.g. there is no word *kabak* ‘pumpkin’ in Benghazi Arabic but there is *bkaywa* ‘pumpkin’, identified by Souag (2013) as a loan from Hausa). Turkish words in Libyan Arabic cited here are from the Benghazi variety, author’s data.

¹⁶For more information on the return of Italian instruction to Libya, see D’Anna (2018).

¹⁷The Italian words in Yoda’s study of Tripoli Judeo-Arabic (Yoda 2005) need to be seen slightly differently than Italian words in non-Jewish dialects, owing to a different history of the Tripolitanian Jewish community with Italy.

Adam Benkato

differently. The primary study devoted to analyzing Italian loans in LA is that of [Abdu \(1988\)](#) who, focusing on the variety of Tripoli, draws up a list of nearly 700 items (a few are misidentified), of which about 50% were recognized by a majority of those surveyed. Some 93% of these are nouns and the remainder are practically all derived from nouns or adjectives, such as *bwōno* ‘well done!’ < *buono* ‘good’ or *faryaz* ‘to go out of order’ < Italian *fuori uso*.¹⁸ Abdu’s study (1988: 248–268) groups Italian loans into some 22 semantic categories, the vast majority of which relate to material culture. Examples of these from the Benghazi variety are *byāmbu* ‘lead’ < *piombo*, *bōskō* ‘zoo’ < *bosco* ‘wood’, *furkayta* ‘fork’ < *forchetta*, *maršabīdi* ‘sidewalk’ < *marciapiede* (author’s data).

As [D’Anna \(2018\)](#) points out, the adaptation of Italian words to Libyan Arabic phonology varies: new phonemes, particularly [v] and [č], sometimes occur but are sometimes adapted to the dialects’ pre-existing phonologies, an indication of “subsidiary phonological borrowing” ([Van Coetsem 1988](#): 98). Of course, the maintenance of new phonemes often depends on speakers continuing to have access to the source language; as this is no longer the case in Libya, Italian borrowings in Libyan Arabic are traversing a different trajectory than French borrowings in other Maghrebi varieties, where only the oldest borrowings have been phonologically integrated.

The overwhelming majority of surviving Turkish and Italian loans in Libyan Arabic are nouns, widely acknowledged to be the most easily-borrowed word class due to their being the least disruptive of the recipient language’s argument structure ([Myers-Scotton 2002](#)), though a few verbs derived dialect-internally do exist. Furthermore, almost all the nouns are cultural borrowings — “lexical content-words that denote an object or concept hitherto unfamiliar to the receiving society, terminology related to institutions that are the property of the neighboring [or colonizing] culture, and so on” ([Matras 2011](#): 210). Cultural borrowings are to be differentiated from core borrowings, the latter being words that more or less duplicate already existing words and which originate in a bilingual codeswitching context. These facts lead us to conclude that Turkish and Italian borrowings in Libyan varieties would be from (1) to (2) on the borrowing scale proposed by Thomason & Kaufman (1988: 78–83). While (1) of the scale involves lexical borrowing of non-basic vocabulary only, (2) includes some function words as well as new phones appearing in those loanwords. Colonial language contact situations are typically ones of recipient-language agentivity, as the number of

¹⁸See [Abdu \(1988: 271\)](#) and [D’Anna \(2018\)](#). Some denominal verbs are cited by Abdu, but more extensive data might reveal several more in use: for example in the variety of Benghazi, I identified *furān* ‘to brake (intransitive)’ < *frayno* ‘brake’ < Italian *freno*, not listed by Abdu.

indigenous people learning the colonial language is many times more than the number of colonizers learning indigenous languages. Without a longer period of sustained bilingualism or language education motivated by continued contact with the metropole, Italian has affected Libyan Arabic to a much less degree than French has Libya's Maghrebi neighbours.

4 Conclusion

The general parameters of the Maghrebi linguistic landscape and contact situations are relatively well understood. However, more documentation of Maghrebi varieties is needed, and more specifically, of those where contact situations – especially with Berber – may have existed. Additionally, further research into the sociolinguistic factors affecting bilingualism in Berber and Arabic, or regarding the intersection of diglossia with bilingualism, will no doubt add to our knowledge of the parameters of contact-induced change more generally. Finally, inter-dialectal contact as well as the gradual rise of national or at least supra-local varieties certainly merits continuing attention.

Further reading

Kossmann (2013b) is the most extensive study so far of Berber–Arabic contact, written from a Berberological point of view but important for Arabists.

Sayahi (2014) studies the intersection of dialects, Standard Arabic, French and Spanish in Tunisia and Morocco.

Souag (2016) summarizes contact in the Saharan region among Arabic, Berber, Hausa, Songhay, Chadic, etc.

Ziamari (2005) is the most up-to-date work discussing code-switching and borrowing strategies between Moroccan Arabic and French.

Adam Benkato

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	MOD	modal
COMP	complementizer	NEG	negative
DEF	definite	PL	plural
DEICT	deictic	PRF	perfect
F	feminine	PRST	presentative
GEN	genitive	REL	relative
IMPF	imperfect	SG	singular
M	masculine		

Acknowledgements

The writing of this article was supported by a grant from the Alexander von Humboldt Stiftung, to whom I am grateful for their support.

References

- Abdu, Hussein Ramadan. 1988. *Italian loanwords in colloquial libyan arabic as spoken in the tripoli region*. Tucson: University of Arizona doctoral dissertation.
- Al-Jallad, Ahmad. 2015. On the voiceless veflex of *š and *ṭ in pre-hilalian maghrebian Arabic. *Zeitschrift für Arabische Linguistik* 62. 88–95.
- Behnstedt, Peter. 2016. The *niktib-niktibu* issue revisited. *Wiener Zeitschrift für die Kunde des Morgenlandes* 106. 21–36.
- Ben Cheneb, Mohammed. 1922. *Mots turks et persans conservés dans le parler algérien*. Algiers: Université d’Alger.
- Cifoletti, Guido. 1994. Italianismi nel dialetto di tunisi. In Dominique Caubet & Martine Vanhove (eds.), *Actes des premières journées internationales de dialectologie arabe de Paris*, 451–458. Paris: Langues’O.
- Colin, Georges S. 1999. *Arabe marocain: Inédits de Georges S. Colin*. Dominique Caubet & Zakia Iraqui-Sinaceur (eds.). Paris: INALCO.
- D’Anna, Luca. 2018. Phonetical and morphological remarks on the adaptation of italian loanwords in libyan arabic. In Stefano Manfredi & Mauro Tosco (eds.), *Arabic in contact*, 172–187. Amsterdam: John Benjamins.

- Gibson, Maik. 2002. Dialect levelling in tunisian Arabic: towards a new standard. In Aleya Rouchdy (ed.), *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, 24–30. London: Routledge.
- Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London: Routledge.
- Heath, Jeffrey. 1989. *From code-switching to borrowing: a case study of Moroccan Arabic*. London: Kegan Paul.
- Kossmann, Maarten. 2010. Parallel system borrowing: parallel morphological systems due to the borrowing of paradigms. *Diachronica* 27(3). 459–487.
- Kossmann, Maarten. 2013a. Borrowing. In Jonathan Owens (ed.), *The Oxford Handbook of Arabic Linguistics*, 349–368. Oxford: Oxford University Press.
- Kossmann, Maarten. 2013b. *The Arabic influence on Northern Berber*. Leiden: Brill.
- Kossmann, Maarten. 2014. On substratum: the history of the focus marker *d* in jijel Arabic (algeria). In Carole de Féral, Maarten Kossmann & Mauro Tosco (eds.), *In and out of africa: languages in question. in honour of robert nicolaï*, vol. 2: *Language contact and language change in Africa*, 127–140. Leuven: Peeters.
- Lafkioui, Mena. 2013. Reinventing negation patterns in moroccan Arabic. In Mena Lafkioui (ed.), *African Arabic: Approaches to dialectology*, 51–93. Berlin: De Gruyter Mouton.
- Lucas, Christopher. 2007. Jespersen’s cycle in Arabic and berber. *Transactions of the Philological Society* 105(3). 398–431.
- Lucas, Christopher. 2010. Negative *-š* in palestinian (and cairene) arabic: present and possible past. *Brill’s Annual of Afroasiatic Languages and Linguistics* 2. 165–201.
- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Lucas, Christopher. 2018. On wilmsen on the development of postverbal negation in dialectal arabic. *Zeitschrift für Arabische Linguistik* 67. 44–70.
- Manfredi, Stefano. 2018. Arabic as a contact language. In Reem Bassiouney & Elabbas Benmamoun (eds.), *The Routledge handbook of Arabic linguistics*, 407–20. London: Routledge.
- Marçais, Philippe. 1956. *Le parler arabe de Djidjelli (Nord Constantinois, Algérie)*. Paris: Maisonneuve.
- Matras, Yaron. 2011. Universals of structural borrowing. In Peter Siemund (ed.), *Linguistic universals and language variation*, 204–233. Berlin: De Gruyter Mouton.

Adam Benkato

- Mion, Giuliano. 2015. Réflexions sur la catégorie des «parlers villageois» en arabe tunisien. *Romano Arabica* 15. 269–279.
- Mourigh, Khalid. 2015. *A grammar of ghomara berber*. Leiden: University of Leiden doctoral dissertation.
- Myers-Scotton, Carol. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Naciri-Azzouz, Amina. 2016. Les variétés arabes de ghomara? *s-saḥəl* vs. *əḡ-ḡbəl* (la côte vs. la montagne). In George Grigore & Gabriel Bițună (eds.), *Arabic varieties: far and wide. proceedings of the 11th international conference of aida, bucharest 2015*, 405–412. Bucharest: Editura Universității din București.
- Pereira, Christophe. 2007. Urbanization and dialect change: the Arabic dialect of tripoli (libya). In Catherine Miller, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.), *Arabic in the city: Studies in dialect contact and language variation*, 77–96. London: Routledge.
- Pereira, Christophe. 2011. Arabic in the north africa region. In Stefan Weninger, Janet C. E. Watson, Geoffrey Khan & Michael P. Streck (eds.), *The Semitic languages: An international handbook*, 954–969. Berlin: De Gruyter Mouton.
- Procházka, Stephan. 2005. The Turkish contribution to the arabic lexicon. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*, 191–202. London: Routledge.
- Procházka, Stephan. 2012. Les mots turcs dans l’arabe marocain. In Alexandrine Barontini, Christophe Pereira, Ángeles Vicente & Karima Ziamari (eds.), *Dynamiques langagières en Arabophonies: Variations, contacts, migrations et créations artistique. Hommage offert à Dominique Caubet par ses élèves et collègues*, 201–222. Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo (IEIOP).
- Sayahi, Lotfi. 2014. *Diglossia and language contact: Language variation and change in North Africa*. Cambridge: Cambridge University Press.
- Sayahi, Lotfi. 2015. Expression of attributive possession in tunisian Arabic: the role of language contact. In Aaron Michael Butts (ed.), *Semitic languages in contact*, 333–347. Leiden: Brill.
- Souag, Lameen. 2013. Sub-Saharan lexical influence in north african arabic and berber. In Mena Lafkioui (ed.), *African Arabic: Approaches to dialectology*, 211–236. Berlin: Mouton de Gruyter.
- Souag, Lameen. 2016. Language contact in the sahara. In Mark Aronoff (ed.), *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press.

4 Maghrebi Arabic

- Souag, Lameen. 2017. Syllable weight and morphophonologically induced resyllabification in maghrebi arabic. In Paul Newman (ed.), *Syllable weight in african languages*, 49–67. Amsterdam: John Benjamins.
- Souag, Lameen. 2018. Arabic–berber–songhay contact and the grammaticalization of ‘thing’. In Stefano Manfredi & Mauro Tosco (eds.), *Arabic in contact*, 54–71. Amsterdam: John Benjamins.
- Taine-Cheikh, Catherine. 2008. Arabe(s) et berbère en contact: le cas mauritanien. In Mena Lafkoui & Vermondo Brugnatelli (eds.), *Berber in contact: Linguistic and sociolinguistic perspectives*, 113–38. Cologne: Rüdiger Köppe.
- Talmoudi, Fathi. 1986. *A morphosemantic study of Romance verbs in the Arabic dialects of Tunis, Sūsa, and Sfax*. Gothenburg: Acta Universitatis Gothoburgensis.
- Thomason, Sarah G. & Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Türkmen, Erkan. 1988. Arapça’nın Libya lehçesindeki türkçe kelimeler [turkish words in the Libyan dialect of arabic]. *Erdem* 4–10. 211–225, 227–243.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Vicente, Ángeles. 2005. *Ceuta: une ville entre deux langues. (une étude sociolinguistique de sa communauté musulmane)*. Paris: L’Harmattan.
- Winford, Donald. 2005. Contact-induced change: classification and processes. *Diachronica* 22(2). 373–427.
- Yoda, Sumikazu. 2005. *The Arabic dialect of the Jews of Tripoli (Libya): Grammar, text and glossary*. Wiesbaden: Harrassowitz.
- Yoda, Sumikazu. 2013. Judeo-arabic, libya, hebrew component in. In Geoffrey Khan (ed.), *Encyclopedia of Hebrew language and linguistics*. Leiden: Brill.
- Ziamari, Karima. 2005. *Le code switching au maroc: l’arabe marocain au contact du français*. Paris: L’Harmattan.

Chapter 5

Moroccan Arabic

Jeffrey Heath

University of Michigan

Morocco, even if the disputed Western Sahara is excluded, is rivaled only by Yemen in its variety of Arabic dialects. Latin/Romance sub- and ad-strata have played crucial roles in this, especially 1) when Arabized Berbers first encountered Romans; 2) during the Muslim and Jewish expulsions from Iberia beginning 1492; and 3) during the colonial and post-colonial periods.

1 History and current state

1.1 History

Moroccan Arabic (MA) initially took shape when Arab-led troops, probably Arabized Berbers from the central Maghreb who spoke a contact Arabic, settled precariously in a triangle of Roman cities/towns consisting of Tangier, Salé, and Volubilis, starting around 698 AD. Mid-seventh-century tombstones from Volubilis, inscribed in Latin, confirm that Roman Christians were present, though in small numbers, when the Arabs arrived. Shortly thereafter, in 710–711, an Arab-led army from Morocco began the conquest of southern Spain, a richer and more secure prize that drew away most of the Arab elite. In Morocco, turnover of the few Arabs and of their Arabized Berber troops was high; they were massacred or put to flight in the Kharijite revolt of 740. The eighth and ninth centuries had perfect conditions for the development of a home-grown Arabic in the Roman triangle in Morocco, and in the emerging Andalus, with a strong Latinate substratum.

The first true Arab city, Fes, was not founded until approximately 798, a century after the first occupation of Morocco, and its population did not bulk up until immigration from Andalus and the central Maghreb began around 817. With a



Jeffrey Heath

cosmopolitan population, and located outside of the old Roman triangle, its Andalusi and non-Andalusi quarters may have maintained their respective dialects for a long time. The remainder of Morocco was occupied by Berber tribes until much later.

During the eleventh century, the Arabian bedouin often called Banu Hilāl entered the central Maghreb in large numbers (cf. Benkato, this volume). They partially bedouinized the Arabic dialects in Tunisia and Algeria, producing hybrid varieties that combined pre- and post-Hilalian features. They also gradually pushed their way south and west across the Sahara, bringing their distinctively bedouin Arabic, known as Ḥassāniyya, into the southern Maghreb, including some oases of southern Morocco proper and the entire Western Sahara. Meanwhile, hybridized Algerian dialects, also reflecting a Berber substratum, were spreading into western Morocco, taking root in new farming villages in the central plains around Fes, and in the younger cities such as Meknes and Marrakesh (Heath 2002).

In 1492, the Catholic Kings abruptly expelled Spanish Jews from Spain, followed by expulsions through 1614 of Muslims from Spain and Portugal (cf. Vicente, this volume). Jewish deportees, whose predominant home language was Judeo-Spanish, flooded into the Jewish quarters (*mellahs*) of Moroccan cities, constituting a new Jewish elite. Muslim deportees, variably speaking Arabic or Romance, arrived in several waves and were more easily assimilated. The Jewish presence in Morocco was strong until 1951, when most Jews left for Israel and other destinations.

Moroccan ports participated in growing Mediterranean and Atlantic maritime activity, associated linguistically with Lingua Franca (cf. Nolan, this volume) and various Romance languages along with Turkish, in the seventeenth and eighteenth centuries. European precolonial penetration into coastal Morocco in the late nineteenth century later expanded during the French and much smaller Spanish protectorates which lasted from 1912 to 1956. Exposure to French increased dramatically in this period.

Also of linguistic relevance is the fact that the Moroccan–Algerian border has been virtually closed for decades, due mainly to political disputes. This has partially sealed off Morocco from the central Maghreb and allowed a specifically Moroccan koiné to flourish.

1.2 Current situation

Of the 33 million Moroccans recorded in a 2014 census, nearly all are fluent L1 or (among the Berber-speaking minority) L2 speakers of some form of MA. More-

over, except in the thinly populated Western Sahara, the once-robust dialectal variation within MA has now been greatly compressed. The MA that one is likely to hear in cafés in Rabat, Fes, Meknes, Marrakesh, Oujda, and even Tangier is the Moroccan koiné, a hybridized variety mixing pre- and post-Hilalian features and showing heavy Berber influence in prosody and vocalism.

Many Berber dialects, commonly (but inaccurately) classified into three languages (Tarifiyt, Tamazight, and Tashelhiyt), are still widely spoken in the mountain ranges and in the Souss valley along the Atlantic coast near Agadir. However, these Berber languages are full of Arabic loans, and they are slowly losing ground to Arabic in all of the cities and large towns.

2 Contact languages

2.1 General

This chapter focuses on contact between MA and European languages. Punic (Phoenecian) had probably died out locally before the Arab conquest, and Greek was a non-factor in spite of nominal Byzantine suzerainty after the fall of the Roman empire. Berber–Arabic contact is covered elsewhere (see Souag, this volume and Benkato, this volume). Diglossic borrowing from literary Arabic would take us far afield; on this, see [Sayahi \(2014\)](#) and [Heath \(1989\)](#).

The hallmark of abrupt language shift is powerful substratal influence in phonology and prosody. Some calquing of grammatical constructions may occur, but this can be difficult to tease apart from morphosyntactic simplification. There may be little or no carryover of core vocabulary and of concrete grammatical morphemes. The profile of language shift contrasts with that of adstratal borrowing during prolonged bilingualism, whose manifestations are mainly lexical, and whose complexities involve the morphological and semantic nativization of foreign-source inflected forms (cf. Manfredi, this volume).

2.2 Late Latin

The best-kept secret about MA is that, unlike the case elsewhere in the Maghreb, its oldest forms originated by language shift (probably rapid) from Late Latin (LL) to a contact Arabic spoken by Berber troops.

There are no written records of colloquial LL of the relevant period, either in North Africa or in Europe, but we can surmise that the LL spoken in the Roman triangle was intermediate between Classical Latin and early Medieval Romance,

Jeffrey Heath

e.g. Medieval Spanish. This implies either five or possibly seven vowel qualities, phonemic stress, no vowel length, and probably some affricates č [tʃ] and ġ [dʒ].

2.3 Medieval Judeo-Spanish

The major injection of medieval Spanish into the Moroccan heartland was the arrival of expelled Spanish Jews in 1492. They joined existing Jewish communities in the large cities, but a cultural divide between the newcomers (*meḡorashim*) and incumbents (*toṣhaviṃ*) quickly emerged. We know from rabbinical responsa that Judeo-Spanish was still spoken in the central cities for two centuries after 1492 (Chetrit 1985). In far northern Morocco, a form of Arabic- and Hebrew-influenced Judeo-Spanish called Hakitia or Haketia remained in vernacular use until the early twentieth century (Benoliel 1977), after which it merged with modern Spanish.

2.4 Modern French and Spanish

Spanish and to some extent Portuguese and Catalan remained contact influences chiefly in ports through the late nineteenth century, when direct Spanish involvement in northern Morocco became more significant. Iberian loanwords figure prominently in the early twentieth-century maritime vocabulary in Brunot (1920). During the Protectorates, French became a major language of education and administration in most of Morocco, especially in the west-to-east Casablanca–Rabat–Fes–Meknes–Taza corridor, while Spanish consolidated its position in the far north. French loanwords during the early Protectorate are in Brunot (1949). MA–French and MA–Spanish bilingualism has increased in the postcolonial period due to media and mass education. English influence is increasing, mainly through tourism, science education, and finance.

3 Contact-induced changes in MA

3.1 Phonology

MA dialects – archaic Pre-Hilalian, hybridized Post-Hilalian, and in the far south the unhybridized Ḥassāniyya – differ sharply in vocalic systems, reflecting their different histories (Heath 2018).

Classical Arabic (CA) had short {ī ā ū} versus long {ī ā ū}, diphthongs {āy āw}, no syncope, and no phonemic stress.

5 Moroccan Arabic

Of the three main types of MA, Ḥassāniyya is closest to CA. It has short vowels limited to closed syllables: {ə ă} with ə < {*ī *ū}, in some dialects (e.g. Mali) also some cases of ŭ. It distinguishes long {ī ā ū} from diphthongs {āy āw}, and has no phonemic stress, but unlike CA it does allow syncope of short vowels (cf. [Taine-Cheikh 1988](#)). Ḥassāniyya shows limited effects of language contact in the phonology of Berber loanwords (cf. [Taine-Cheikh 1997](#)).

By contrast, the koiné and some other hybrids reduce all three short vowels to just one short vowel ə with various allophones, contrasting with full vowels {i a u}. The hybrid dialects monophthongize {*āy *āw} to merge with {i u}. The rounding of original short *ū often survives next to a velar/uvular consonant, even after syncope (which is productive), suggesting an ongoing feature transfer that, if and when fully implemented, would result in underlying labiovelars {k^w g^w q^w ɣ^w ʕ^w} next to ə (which becomes phonetic [ʊ]) or before a consonant. Again there is no phonemic stress. This is a Berber-like system, reflecting deep long-term substratal/adstratal contact.

A more archaic Berber-like system, still preserving at least the opposition of short *ī ~ *ā versus *ū and likely at least some diphthongs, was brought to Morocco by the early Arabized Berber troops. There it was overlaid on an LL substratum that had five to seven vowel qualities, phonemic stress, no syncope, and no vowel length. The resulting Pre-Hilalian MA has: three regular vowels {i a u}, a subset of which (the original short vowels) syncopate in weak metrical positions; phonemic stress; and a schwa vowel ə confined to posttonic final closed syllables. The leveling of vowel length distinctions, and the re-splitting of the previously merged *ī ~ *ā into i and a based on consonantal environment, were disruptive to the morphology (see §3.2). Both the leveling, and the new phonemic stress, were shared with speakers of early Andalusī Arabic, which had a similar LL substratum and whose first invaders came from Morocco. This points to an original dialect area in the eighth and ninth centuries, including coastal Andalus and at least the Tangier–Salé axis in Morocco (after Volubilis was abandoned in favor of Fes), differing significantly from even Pre-Hilalian central Maghrebi dialects, which likely never had major LL substratal effects.

The differences among MA dialect types can be illustrated by forms of ‘big’ (Table 1). The suggested proto-forms are close to CA but show some adjustments to short *ī and *ū. Acute accent marks stress in Pre-Hilalian. Observe especially that the two homophonous Pre-Hilalian *kbīr* forms behave differently when a vowel-initial suffix is added. The morphological consequences of length merger in Pre-Hilalian are considered below. Emphatic /ɾ/ is phonemically distinct from plain /r/ in all varieties.

Jeffrey Heath

Table 1: The word-family ‘big’ in MA dialect types

Gloss	Proto	Pre-Hilalian	Hybrid	Ḥassāniyya
‘big’ (SG.M)	*kābīr	<i>kbír</i>	<i>kbir</i>	<i>kbīr</i>
‘big’ (SG.F)	*kābīr-a	<i>kbír-a</i>	<i>kbir-a</i>	<i>kbīr-a</i>
‘he got big’	*kābīr	<i>kbír</i>	<i>kbər</i>	<i>kbər</i>
‘she got big’	*kābīr-at	<i>kíbr-at</i>	<i>kbr-at</i>	<i>kəbr-at</i>
‘bigger’	*ākḃār	<i>kbār</i>	<i>kbər</i>	(ǎ)kbār
‘big (PL)’	*kūbār	<i>kbaṛ-ín</i>	<i>k^wbaṛ</i>	<i>kbār</i>

Later adstratal borrowings from Spanish and French, as well as from CA, predictably required adjustments to MA phonology. The most disruptive changes affected French borrowings into MA (our data are best for the hybrid koiné). The rich array of French vowel qualities had to be squeezed into three MA qualities. French $\{i \ddot{u} e \epsilon\}$ merge as MA *i*. French $\{u o \circ \ae\}$ merge as MA *u*. French *a* becomes MA *a*. This compression has had considerable morphological consequences (see §3.2 below).

The main contribution of Romance to MA consonantism is the affricate \check{c} [tʃ]. In the current koiné, this is present as a phoneme (if at all) in the loanword *lččina* ~ *ltšina* ‘orange (fruit)’ < Spanish *la China*, as brought out in the diminutive which breaks up the $\check{c}\check{c}$ cluster, hence *lčičin* ~ *ltišin* and further variants (Heath 1999). Archaic northern dialects have more examples of \check{c} , and these dialects pronounce geminated \check{z} as affricate \check{g} [dʒ].

3.2 Morphology

Direct borrowing of bound function morphemes is rare in MA as in other languages. A notorious exception is *ta*...-*t* in abstract nouns of profession, from the Berber feminine singular, likely extrapolated from specific Berber borrowings like *ta-šaffar-t* ‘thief’.

Another glaring exception is the set of D-possessives: *d* (archaic *dī*) before nouns, *dyal*- (Pre-Hilalian *dyál*-) primarily before pronominal suffixes (e.g. *dyal-i* ‘mine’, *dyal-u* ‘his’). The obvious etymology (Latin *dē* → LL **de* or unstressed **dī*) presents no phonological or semantic difficulties, but it was rejected by a century of Maghrebi Arabists, who favored various far-fetched Arabic-internal etymologies. However, an LL source is also indicated by its dialectal distribution: Pre-Hilalian MA, regional colloquial Andalusi Arabic, and certain coastal

enclaves in Algeria that were likely settled by Andalusí merchants. The mysterious prepronominal variant *dyál-* was generalized from LL **di él(l)u* ‘his’ and LL **di él(l)a* ‘hers’, which are near-exact matches to the still extant Pre-Hilalian *dyál-u* ‘his’ and *dyál-a* ‘hers’. The motivation for this admittedly unusual morphemic borrowing was the need for a new possessive morpheme as Arabic dialects gradually abandoned the compound-like CA “construct” possessive (Heath 2015). The fact that possessive morphemes are not immune from borrowing is also shown by possessed forms of certain kin terms, with a Berber nasal suffix, before nominal possessors in hybrid dialects, as in (koiné) *b̥ba-yn ḥamid* ‘Hamid’s father’, cf. *b̥ba* ‘father’.

Verbs as well as nouns are readily borrowed from Romance languages into MA. This raises the question of which Romance inflected form is borrowed, and what value it is assigned to within the MA tense system, which groups 1st/2nd persons versus 3rd person subject splits in the perfect of some verb types. Most Spanish verb borrowings look like Spanish infinitives, e.g. *f̣rinar* ‘to brake’ (< *frenar*), but more likely reflect a cluster of forms based on this stem shape in Spanish itself. In addition to the infinitive, this set also includes future *frenar-é*, conditional *frenar-ía*, and forms with *d* instead of *r*, namely participle *frenado* and imperative plural *frenad*. Consonant-final borrowed verbs like *f̣rinar* behave like native MA quadriliteral verbs, and have identical perfect and imperfect forms.

By contrast, French verbs are regularly borrowed as weak (i.e. vowel-final) verbs, with imperfect and 1st/2nd perfect *i*, versus third-person perfect *a*. An example is ‘declare’: imperfect -*ḍiklari* matching perfect 1st/2nd *ḍiklari-*, versus 3rd *ḍiklara(-)*. The likely crosslinguistic bridge is the conspicuous cluster of French forms ending in orthographic -*er* (infinitive), -*ez* (2PL subject), -*ais/-ait/-aient* (imperfect), and -*é(e)(s)* (participle). All of these are phonetic [e] or [ɛ] and therefore merge as MA *i*, interpretable in MA as the imperfect and 1st/2nd perfect of weak verbs. The marked third-person perfect with final *a* is then easily formed by analogy (cf. Lucas & Čéplö, this volume: §4.2. for a parallel development in Maltese).

The merger of vowel length in Pre-Hilalian MA set off a chain reaction of morphophonological restructurings, most notably in the verbal system. The CA three-way vocalic opposition of hollow verbs, e.g. for ‘to be’ imperfect *kūn-*, pre-consonantal perfect *kūn-*, and prevocalic (or word-final) perfect *kān-*, is largely preserved in hybrid and Post-Hilalian dialects. By contrast, in Pre-Hilalian MA, after the momentous vowel-length merger, the hollow paradigm was reorganized into a binary opposition of *kún* (imperfect and 1st/2nd perfect) versus *kán* (3rd perfect). This paradigmatic reorganization, which makes no sense semantically and is apparently unique to Pre-Hilalian MA, then spread analogically to other

Jeffrey Heath

verb types, including strong trilaterals that have three consonants and no long vowels, e.g. ‘enter’: imperfect *-thul* matching 1st/2nd perfect *thul*-, but 3rd perfect *thal*.

3.3 Syntax

Before reaching Morocco, spoken Arabic had prepositions, possessum–possessor, and DEF–N–ADJ order within NPs, preverbal negation (cf. Lucas, this volume) and complementizers, a perfect/imperfect split in verbs, and pronominal-subject agreement on verbs (expressed, in part, by suffixes). Romance languages like Spanish, and presumably eighth-century LL, were already close to this profile, so opportunities for syntactic influence were limited. Some minor French complementizers are common in educated MA, as in *au lieu d’ igulu* ... ‘instead of them saying’, from French *au lieu de* ‘instead of’ plus MA *igulu* ‘they say’.

3.4 Lexicon

While the LL substratum had a profound effect on early MA phonology and morphophonemics, and also left behind a morphemic souvenir in the form of D-possessives, not a single basic LL lexical item can be shown to have been preserved in any archaic MA dialect. The most promising candidate for such a retention is dialectal MA *qḅtal* and variants ‘elbow’. The likely etymon is LL **cubitellu* (later LL **kubtellu*), diminutive of Latin *cubitu(s)* ‘elbow’, cf. modern Spanish *codillo*. The other possibility, less straightforward semantically, is a reflex of the related adjective, Latin *cubitāle*, cf. modern Spanish *codal*. In Morocco, *qḅtal* ‘elbow’ survives in several Judeo-Arabic dialects. For Muslims, it was recorded in an unspecified location in the unpublished fichier of colonial-period linguist Georges Colin (Iraqi Sinaceur 1993: 1525; de Prémare 1998: 224), and by me in the 1980s in archaic varieties of the Fes–Sefrou area. *qḅtal* is completely unknown to the great majority of Moroccan Muslims. Preservation of *b* shows that *qḅtal* is not a recent borrowing from any form based on modern Spanish *codo*. The *b* was still present in (very) Old Spanish *cobdo*, its diminutive *cobdillo*, and *cobdal*. “*Cubtill*” ‘elbow’ is recorded for late Andalusī Arabic (Corriente 1997: 412; Dozy 1967: 302). The geographic and communal distribution of *qḅtal*, especially among Muslims, suggests that it was introduced into Morocco by late Medieval Jewish refugees.

There are, however, hundreds of well-established Spanish loanwords, especially in northern Morocco. There, Spanish is ubiquitous in schools and broadcast media, Spanish tourists are common, and many Moroccans serve as day-laborers

5 Moroccan Arabic

in Spanish enclaves Ceuta and Melilla. While Spanish got a precolonial head-start, French has long since overtaken it in the rest of Morocco. Of special interest are cases where an original Spanish borrowing was later gallicized, sometimes only in part. Examples are MA *antiris* ‘(monetary) interest’, a hybrid of Spanish *interés* and French *intérêt*, and MA *grabaṭa* ‘necktie’ from Spanish *corbata* and French *cravate*. Nonsynonymous mergers also occur, as with *garṣun*, attested both as ‘waiter’ (French *garçon*) and ‘underpants’ (Spanish *calzón*). ‘To sign’ is now usually *-siji/sija* or *-sini/sina* (< French *signer*), but an obsolescent Judeo-Arabic variant *sijaṛ* with (pseudo-)Spanish infinitival ending is attested. Since the Spanish synonym is the unrelated *firmar*, MA *sijaṛ* must have been formed by applying a borrowing routine “add *-aṛ* to the stem” to French stems, probably early in the colonial period when still-abundant Spanish borrowings were being replaced or hybridized under the influence of the newly dominant French.

The process is now coming full circle, as English influence expands. The weak verb alternation of final *a/i* is productive for verbs borrowed from French, as noted above (cf. again the close parallels in Maltese; Lucas & Čěplö, this volume: §4.2). A borrowing routine “add final *a/i* to the stem” extrapolated from French/MA pairs, is now extended to English, where it has no basis in English inflectional paradigms. Examples are the comical *ka-y-spiki mzyan* ‘he speaks (English) well’, and junkie slang like *tt-ṣṭuna* ‘he got stoned’ (participle *m-ṣṭuni* ‘stoned’).

And then there are the many playful translinguistic inventions, concocted among groups of men sitting in cafés, sipping mint tea or smoking... whatever. Nearly all such inventions are ephemeral, but a few have caught on (Heath 1987). Consider the fairly common koiné noun *ḥwadri* ‘pal, buddy’. Unbeknownst to those who now use it, it must have arisen via two successive transformations. First, Spanish *padre* and *madre* were playfully combined with the CCaCCi template for denominal occupational derivatives, as though derived from MA *bḥa ~ bu* ‘father’ and MA *mḥa^(w)* ‘mother’. Templatic CCa... is realized as Cwa... when based on a CV... input, as in *ṣwabni* ‘seller of soap’ (< *ṣabun*). Combining CCaCCi with *padre* and *madre* produces the slang terms (attested but rare) *pḥwadri* and *mḥwadri*. The final and most ingenious step was to combine the sub-template Cwadr-i, emergent from these ‘father/mother’ forms, to *ḥa- ~ hu-* ‘brother’, outputting *ḥwadri*, which then acquires the same ‘buddy’ sense as American English *bro*.

Jeffrey Heath

4 Conclusion and prospects

The broad outlines of historical language contact in Morocco are becoming reasonably clear. The most urgent need is for more material and analysis of Moroccan Judeo-Arabic (MJA), in forms accessible to international audiences. Ideally we would want to tease apart the original LL influence on Pre-Hilalian MJA, as preserved by the *toshavim*, from the medieval Judeo-Spanish brought to Morocco in 1492 by the *megorashim*.

Significant Moroccan Arab and Berber expat communities exist in France, the Netherlands, Belgium, Germany, Switzerland, and Spain. These *vacanciers* return to Morocco in large numbers during summer vacations and on Muslim holy days. There are opportunities to study them both in Europe (Nortier 1990) and in their interactions with other Moroccans.

Another promising topic for investigation is a semi-pidginized form of MA used by monolingual maids in large cities as a kind of foreigner talk to their expat French employers.

Further reading

Heath (1989) is a study of lexical and phrasal borrowing/codeswitching from European languages and from standard Arabic in Moroccan Arabic.

Nortier (1990) examines language contact phenomena among Moroccans in the Netherlands.

Sayahi (2014) is a regional study of Arabic sociolinguistics and language contact from Spain through Morocco to Tunisia.

Acknowledgements

Fieldwork in Morocco was supported by grants from the National Science Foundation (especially BNS 79-04779 in 1979-81) and by a Fulbright research fellowship in 1986. For support while working on MA material in the 1980s, thanks also to the National Endowment for the Humanities, the Deutscher Akademischer Austauschdienst, the Alexander von Humboldt Stiftung, and the Hebrew University of Jerusalem.

Abbreviations

CA	Classical Arabic
F	feminine
JMA	Jewish Moroccan Arabic
L2	second language
LL	Late Latin
MA	Moroccan Arabic
M	masculine
PL	plural
SG	singular

References

- Benoliel, José. 1977. *Dialecto judeo-hispano-marroquí o hakitia*. Madrid: Varona.
- Brunot, Louis. 1920. *Notes lexicologiques sur le vocabulaire maritime de Rabat & Salé*. Paris: Ernest Leroux.
- Brunot, Louis. 1949. Emprunts dialectaux arabes à la langue française depuis 1912. *Hespéris* 36. 347–430.
- Chetrit, Joseph. 1985. Judeo-Arabic and Judeo-Spanish in Morocco and their sociolinguistic interaction. In Joshua Fishman (ed.), *Readings in the sociology of Jewish languages*, 261–279. Leiden: Brill.
- Corriente, Federico. 1997. *A dictionary of Andalusí Arabic*. Leiden: Brill.
- Dozy, Reinhart. 1967. *Supplément aux dictionnaires arabes*. Third edition. Vol. 2. Leiden/Paris: Brill & Maisonneuve et Larose.
- Heath, Jeffrey. 1987. *Hasta la mujerra!* and other instances of playful language mixing in Morocco. *Mediterranean Language Review* 2. 113–116.
- Heath, Jeffrey. 1999. Sino-moroccan citrus: Borrowing as a natural scientific experiment. In Lutz Edzard & Mohammed Nekroumi (eds.), *Tradition and innovation: Norm and deviation in Arabic and Semitic Linguistics*, 168–176. Wiesbaden: Harrassowitz.
- Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London: Routledge.
- Heath, Jeffrey. 2015. D-possessives and the origins of moroccan Arabic. *Diachronica* 32(1). 1–33.
- Heath, Jeffrey. 2018. Vowel-length merger and its consequences in archaic moroccan Arabic. *Zeitschrift für Arabische Linguistik* 6. 12–43.
- Heath, Jeffrey. 1989. *From code-switching to borrowing: a case study of Moroccan Arabic*. London: Kegan Paul.

Jeffrey Heath

- Iraqi Sinaceur, Zakia. 1993. *Le dictionnaire Colin d'arabe dialectal marocain*. Vol. 6. Rabat: Institut d'Etudes et de Recherches pour l'Arabisation.
- Nortier, Jacomine. 1990. *Dutch–Moroccan Arabic code switching*. Dordrecht: Foris.
- de Prémare, André-Louis. 1998. *Dictionnaire arabe-français*. Vol. 10. Paris: l'Harmattan.
- Sayahi, Lotfi. 2014. *Diglossia and language contact: Language variation and change in North Africa*. Cambridge: Cambridge University Press.
- Taine-Cheikh, Catherine. 1988. Métathèse, syncope, épenthèse: a propos de la structure prosodique du hassaniyya. *Bulletin de la Société de Linguistique de Paris* 83(1). 213–252.
- Taine-Cheikh, Catherine. 1997. Les emprunts au berbère zénaga: Un sous-système vocalique du ḥassāniyya. *Matériaux arabes et sudarabiques (GELLAS)* 8. 93–142.

Chapter 6

Maltese

Christopher Lucas

SOAS University of London

Slavomír Čéplö

Institute of Oriental Studies, Slovak Academy of Sciences / IMAFO Abteilung
Byzanzforschung, Österreichische Akademie der Wissenschaften

This chapter presents an overview of the most prominent contact-induced developments in the history of Maltese, a language which is genetically a variety of Arabic, but which has undergone significant changes, largely as a result of lengthy contact with Sicilian, Italian, and English. We first address the precise affiliation of Maltese and the nature of the historical and ongoing contact situations, before detailing relevant developments in the realms of phonology, inflectional and derivational morphology, syntax and lexicon.

1 Maltese and Arabic

From a historical point of view, Maltese is a variety of spoken Arabic, albeit one that has undergone far-reaching changes as a result of sustained and intensive contact with Italo-Romance varieties, and more recently also with English. This is a fact about which there is no controversy among contemporary linguists. It should be noted, however, that a mix of social, cultural, historical, political, and indeed linguistic factors has led to a situation in which many Maltese people today view their language as Semitic, but not a type of Arabic. Since we are concerned here only with the historical perspective, we will not dwell on the vexed question of whether or not contemporary Maltese should be classified as an “Arabic dialect”.¹ Suffice it to say that the idea, first popularized by **de Soldanis**

¹Note that Maltese itself has a number of different dialects, one of which – that of the major towns, and the variety used in media, literature and administration – is referred to as stan-



(1750) and Vassalli (1791), that Maltese is a variety of Phoenician or Punic, has been shown since at least since Gesenius (1810) and de Sacy (1829) to be entirely without merit.

Since the Phoenicians and then the Carthaginians occupied Malta for much of the first millennium BCE, followed by Roman and Byzantine occupation for much of the first millennium CE, it would seem *prima facie* likely that elements of the languages of these occupiers would survive into contemporary Maltese. Brincat (1995) shows, however, based on the account of al-Ĥimyarī, that Malta was to all intents and purposes uninhabited in the period between its conquest by the Arabs in 870 CE and the first concerted efforts at colonization by Arabic-speaking Muslims in 1048–9 CE. It is for this reason that the Semitic component of Maltese phonology, morphology, syntax and lexicon is Arabic and Arabic only (see also Grech 1961).

As for the provenance of the Arabic component of contemporary Maltese, there is no doubt that the most important source is a variety of Maghrebi (Western) Arabic. This is evident from grammatical features such as: the pan-Maghrebi extension to the singular of the first-person *n-* prefix of the imperfect verbal paradigm (see Table 1); the loss of a gender distinction in the second person singular, in pronouns and both perfect and imperfect verbs, as in urban Tunisian Arabic varieties (Gibson 2011); variable rearticulation of the definite article on postnominal adjectives in definite noun phrases, as in (1) (cf. Gatt 2018), found also in Casablanca Arabic (Harrell 2004: 205); and the *-il* suffix of the numerals ‘eleven’ to ‘nineteen’ in determiner use, as in (2), which also occurs in the Arabic dialects of Casablanca (Caubet 2011) and Tlemcen (Taine-Cheikh 2011).²

Table 1: First person imperfect ‘write’ in Eastern and Western Arabic

	Eastern		Western	
	Classical Arabic	Baghdad Arabic	Casablanca Arabic	Maltese
Singular	<i>ʔaktub</i>	<i>aktib</i>	<i>nəktəb</i>	<i>nikteb</i>
Plural	<i>naktub</i>	<i>niktib</i>	<i>nkətbu</i>	<i>niktbu</i>

dard Maltese. Except where specified, this chapter deals exclusively with the standard variety of Maltese.

²Unless otherwise specified, all numbered examples present data from Maltese. All Maltese examples in this chapter are rendered using standard Maltese orthography.

- (1) il-kelb (l-)abjad
DEF-dog (DEF)-white
'the white dog'
- (2) it-tnax-il appostlu
DEF-twelve-DEF apostle
'the twelve apostles'

Narrowing matters down further, Zammit's (2014) study of lexicon shared between Maltese and the Arabic dialect of Sfax offers yet more support (see also Vanhove 1998) for the geographically unsurprising conclusion that Maltese is more closely related to the traditional (so-called pre-Hilalian; see Benkato, this volume) urban Tunisian dialects than to any other extant Arabic variety. This is not to suggest, however, that the Arabic component of Maltese resembles these dialects in all respects. Borg (1996) lists a number of areas in which Maltese accords more closely with Levantine Arabic dialects than with those of the Maghreb. But the social and political history of Malta after the end of direct Arab rule in 1127 CE is such that most or all of these similarities should be understood as the failure of Maltese to participate in innovations that later spread through the mainland Maghrebi varieties, and not as evidence of influence of Eastern Arabic on the formation of Maltese.

2 Contact with Italo-Romance and English

2.1 Italo-Romance

A comprehensive history of immigration to Malta in the medieval period is yet to be written (if indeed such a history is possible at all, given the apparently scarce documentary evidence). It is therefore impossible to give precise details of the sociolinguistic conditions under which the Arabic variety spoken in Malta came into contact with varieties of Italo-Romance in the course of the second millennium. We can, however, sketch the broad outlines of this process, and make some reasonable inferences.

The Arabic-speaking settlers who colonized Malta in 1048–9 CE can be assumed to have come from either Sicily or southern Italy or both (Brincat 1995: 22), but in any case it seems likely that at least some of these came speaking a variety of Sicilian in addition to Arabic. Even after Malta was brought under Norman control in 1127 CE by Roger II of Sicily, and went on to be part of the Kingdom of Sicily, there does not, however, seem to have been really large-scale

immigration of non-Arabic speakers to Malta at any point, a fact which is of course consistent with the survival of the Maltese language until today. Unsurprisingly from a geographical and political perspective, what immigration there was appears to have come overwhelmingly from Sicily and southern Italy, with lesser numbers coming also from Spain (Ballou 1893: 134, 289; Blouet 1967: 43–46; Fiorini 1986; Goodwin 2002: 26–32).

Comprising mostly soldiers, craftsmen and churchmen of various types, it would appear that this immigration was disproportionately male. In addition to families in which the only language spoken was Maltese, there must, therefore, have been significant numbers of families in medieval Malta in which the father spoke only Sicilian natively and the mother spoke only Maltese natively, with communication necessarily involving second-language speech by one or both parents. Children of such families would therefore have been exposed minimally to native and non-native Maltese speech and native Sicilian speech.

From the perspective of Van Coetsem's (1988; 2000) framework for understanding contact-induced change, therefore, it seems highly likely that transfer from Sicilian to Maltese occurred both through imposition under source-language agentivity (by L1 Sicilian speakers) and borrowing under recipient-language agentivity (by L1 Maltese speakers).

There is no doubt that, alongside Sicilian, (Tuscan) Italian had an important place in Maltese life over many centuries, starting at the latest in 1530, when it became the official language of government under the regime of the Knights of Malta. But as Comrie & Spagnol (2016: 316) point out, Italian did not gain a foothold at the expense of Sicilian among bilingual Maltese until the later eighteenth century, and given its social function as a vehicle for government, education and high culture, rather than the native language of a significant proportion of ordinary Maltese, it is reasonable to say that transfer from Italian will have been mediated predominantly by borrowing under recipient-language agentivity.

2.2 English

Starting in 1800, when Malta became a protectorate of the British Empire, English gradually began to supplant Italian as the language of government, education and high culture, being joined in that role by the Maltese language itself only in the last few decades. English is now widely spoken in Malta: according to 2011 census data (*Census of Population and Housing 2011: Final Report 2014*: 149), 94.6% of the population of Malta reported speaking Maltese “well” or “average[ly]”, while 82.1% reported the same for English. English is a native language for only a very

small percentage of Maltese residents, however: Sciriha & Vassallo (2006) put the figure at 2%. As with Italian, then, transfer from English to Maltese will overwhelmingly have occurred through borrowing under recipient-language agentivity. With the Maltese variety of English, the reverse is true of course: here the transfer from English to Maltese will have been almost exclusively imposition under source-language agentivity by native speakers of Maltese, resulting in such hallmark features of Maltese English as word-final obstruent devoicing (cf. §3.1.1.2 below), and the use of *but* in clause-final position (Lucas 2015: 527).

Given that transfer from English was and is restricted to borrowing in Van Coetsem's sense, while the more extensive and long-lasting contact with Sicilian will have involved both borrowing and imposition, it is not surprising that a picture will emerge in the following sections whereby Italo-Romance dominates as a source of contact-induced changes across all linguistic domains, with English playing a much more modest role, largely restricted to lexicon and associated inflectional morphology.

3 Contact-induced changes

3.1 Phonology³

3.1.1 Consonants

3.1.1.1 Additions to the native phonemic inventory

One of the most salient, and uncontroversially contact-induced, innovations in Maltese phonology is the addition of at least five (arguably seven) consonant phonemes. This came about through the transfer (presumably borrowing) of Italo-Romance and English lexical items without subsequent adaptation to the original native inventory (compare, e.g., Maltese *pulizija* with unadapted initial [p] and Cairene Arabic *bulīs* 'police'). The five uncontroversial additions are /p/, /v/, /ts/, /tʃ/ and /g/ (orthographically: <p>, <v>, <z>, <ċ> and <g>; see Table 2), as in *evaporazzjoni* 'evaporation' and *graniċ* 'crab'. One can also make a case for an innovative borrowed phoneme /dʒ/. There are no minimal pairs demonstrating a phonemic distinction between /dʒ/ and /ts/ (and both are represented by <z> in the orthography), but Borg & Azzopardi-Alexander (1997: 301) point out that /dʒ/ occurs in environments not requiring a voiced obstruent, as in *gazzetta* /gɛdʒ:ɛt:ɐ/ 'newspaper'. More marginal is /ʒ/, which Mifsud (2011) and Borg & Azzopardi-Alexander (1997: 303) point out can be found in recent loanwords

³For useful overviews of the phonology of Maltese, see Borg (1997) and Cohen (1966; 1970).

from English, such as *televixin* ‘television’ and *bex* ‘beige’, though whether all speakers voice the <x> in these items is uncertain.

Table 2: Inventory of consonants. Symbols are Maltese orthography.

	Labial	Alveolar	Post-alveolar	Palatal	Velar	Laryngeal
Plosive	p b	t d			k g	q
Affricate		z	č ġ			
Fricative	f v	s z	x			ħ
Nasal	m	n				
Trill		r				
Lateral		l				
Approximant	w			j		

Proto-Semitic *g, represented as <ġ> in Arabic script, and usually rendered [dʒ] when standard Arabic is spoken, is reflected as /dʒ/ (orthographic <ġ>) in Maltese. This appears to be a retention of the original Maghrebi realization of this phoneme, other Maghrebi varieties having in general deaffricated it to /ʒ/ (cf. [Heath 2002](#): 136). Unlike some other Maghrebi varieties, however, the Maltese reflex of <ġ> does not become /g/ before sibilants (cp. Maltese *ġewż* vs. Casablanca *gūz* ‘walnuts’).⁴ Similarly, Proto-Semitic *q (on which more below), is never reflected as /g/ (orthographic <g>) in Maltese (cf. [Vanhove 1998](#): 99), meaning that the presence of /g/ in the Maltese phonemic inventory is certainly due to its occurrence in numerous lexical borrowings. The majority of these are from Italo-Romance (e.g. *gwerri* ‘war’), but some are from Berber (e.g. *gendus* ‘calf’ < Berber *agenduz*, [Naït-Zerrad 2002](#): 827), suggesting that /g/ as an independent phoneme has been present in Maltese since the earliest days of Arabic speech on the islands.⁵

⁴An exception is *gżira* ‘island’ < Arabic *ġazīra*, perhaps to be explained by direct contiguity with the sibilant.

⁵There are also some sporadic examples of /g/ < *k in Arabic roots, e.g. *gideb* ‘to lie’. See Cohen (1966: 14–15) for further details.

3.1.1.2 Losses, mergers and shifts

Alongside these additions, the Maltese consonant phoneme inventory has also witnessed a number of losses and mergers. Clearly it is not possible to establish with certainty whether or not these changes were due to contact, but various considerations make it reasonable to assume that contact at least accelerated these changes. For example, the inherited emphatic (pharyngealized/uvularized) consonants – *ṣ, *ṭ, *ḍ, and *ẓ – have all merged with their non-emphatic counterparts, as in *ṣḥab* /ṣħə:b/ ‘clouds’ < *saḥāb*, and also ‘companions’ < *ʔaṣḥāb*. Note in this connection that among other Arabic varieties, it is only a handful of those most strongly affected by contact (such as pidgins and creoles, as well as Cypriot Maronite Arabic; see Avram, this volume; Walter, this volume) that have merged the emphatic consonants in this way. This suggests that non-native acquisition of Maltese by Italo-Romance speakers precipitated this change (i.e. that it involves source-language agentivity in Van Coetsem’s 1988; 2000 terms).

In addition to the loss of the emphatic consonants, Maltese has undergone significant losses and mergers among the velar and laryngeal phonemes.

Perhaps most saliently, an earlier version of what is today standard Maltese merged and then lost the voiced uvular/velar fricative *ɣ and the voiced pharyngeal fricative *ʕ. In Maltese’s rather etymologizing orthography, these historic phonemes are given the digraph symbol <gh>. In general, this symbol either has no phonetic correlate, as in *ghajn* /ɐm/ ‘eye, spring’ and *ghasfur* /ɐ:sʃu:r/ ‘bird’, or otherwise corresponds to the lengthening of a vowel in morphological patterns where the vowel would ordinarily be short, as in the stem I *CaCeC* verb *ghamel* /ʔe:mɛl/ ‘to do’. That the two original phonemes first merged and were then lost in standard Maltese can be inferred from the behaviour of <gh> + <h> sequences. These are realised as /ħ:/ in roots where <gh> reflects *ʕ (e.g. *semagħ-ha* /sɛ'mɛħ:ɐ/ hear.PRF.3SG.M-3SG.F, ‘he heard it’), where other Arabic varieties behave similarly (cf. Woidich 2006: 18), but also, unlike other Arabic varieties, in roots where <gh> reflects *ɣ (e.g. *ferragħ-ha* /fɛr'rɛħ:ɐ/ pour.PRF.3SG.M-3SG.F, ‘he poured it out’ < Arabic *faray* ‘to empty’). This merger and subsequent loss did not take place in all varieties of Maltese. To this day, there are apparently speakers of dialectal Maltese whose speech preserves both *ɣ as a velar fricative, and *ʕ as a pharyngeal fricative (Klimiuk 2018). The fact that the merger and loss of these two phonemes is more advanced in the standard language of the major conurbations and less so in the dialects of more isolated villages suggests that contact-induced change played an important role here, with non-native speakers of Maltese presumably being the principal agents of change.

Arguably the most interesting set of mergers and losses concerns the voice-

less fricatives, which represent a case of considerable phonemic reorganization despite relatively little change at the phonetic level. The phonemic changes in this domain are as follows. First, *h, while maintained in the orthography (as <h>), has merged with /h/ in codas (e.g. *ikrah* /ik'rɛh/ 'ugly') and sporadically in onsets (e.g. *naħaq* /nɛħɛʔ/ < *nahaq* 'to bray (of donkeys)'), and is otherwise lost altogether (e.g. *hemm* /ɛm:/ 'there'). The Maltese phoneme /ħ/ thus represents the continuation of the voiceless pharyngeal fricative *ħ, as well as the partial merger of *h. Moreover, original *ħ, the voiceless uvular/velar fricative, has also merged with /h/, as in *ħajt* 'thread' < *ħayt*, and also 'wall' < *ħāyit*. Strikingly, however, the single Maltese phoneme /ħ/ exhibits considerable inter- and intra-speaker variation in its precise realization, such that glottal, pharyngeal, and velar/uvular voiceless fricative realizations may commonly be heard (Borg & Azzopardi-Alexander 1997: 301), and it is in this sense there has been little phonetic change despite the considerable phonological reorganization.

Like the loss of the emphatic consonants, the loss or merger of *h, as well as one or more of the pharyngeal and velar/uvular fricatives, is restricted to a handful of Arabic varieties that have been very strongly affected by contact (see, e.g., Walter, this volume). As such, these changes too are suggestive of imposition by non-native speakers lacking these sounds in their native phonemic inventory (as was the case for speakers of the Romance varieties with which Maltese has had the most intense contact, cf. Loporcaro 2011: 141–142). On the other hand, the preservation of the glottal and pharyngeal fricatives as allophones of /ħ/ complicates this picture, such that the role of contact in bringing about these particular changes must remain uncertain for now.

It is similarly hard to diagnose the causes of the shift of *q to glottal stop (nevertheless written as <q> in Maltese orthography) and the stopping of the interdental fricatives *θ and *ð. In both cases, however, we can at least rule out with confidence any suggestion that these are ancient changes that predate the arrival of Arabic in Malta, or are historically connected to similar realizations in the Arabic dialects of urban centres in the Maghreb, Egypt, and the Levant. Written records of earlier Maltese clearly show that a dorsal realization of *q, as well as the interdental fricative realization of *θ and *ð, survived until at least the late eighteenth century (Avram 2012; 2014). It is at least plausible, therefore, that contact with Italo-Romance played a role in these changes too, but firm evidence on this point is so far lacking.

Finally, a well-known feature of contemporary Maltese (and Maltese English) phonology is the devoicing of word-final obstruents, as in *ħadd* [ħɛt:] 'nobody'. Avram (2017) shows that devoicing gradually diffused across the Maltese lexicon

over the course of about two centuries from the late sixteenth century onwards, and he makes a strong case that the initial trigger for this development was imposition by native speakers of Sicilian and Italian, since word-final obstruent devoicing has been shown by various studies (e.g. [Flege et al. 1995](#)) to be a frequent feature of the L2 speech of L1 speakers of Romance languages.

3.1.2 Vowels

Maltese has a much richer vowel phoneme inventory than typical Maghrebi Arabic dialects, with, among the monophthongs, five short-vowel qualities /ɪ, ɛ, ɐ, ɔ, ʊ/ (orthographic <i, e, a, o, u>), and six long-vowel qualities /i:, ɪ:, ɛ:, ɐ:, ɔ:, u:/ (orthographic <i, ie, e, a, o, u>), as well as seven distinct diphthongs (with a number of different orthographies – see [Borg & Azzopardi-Alexander 1997](#): 299 for details): /ɪʊ, ɛɪ, ɛʊ, ɐɪ, ɐʊ, ɔɪ, ɔʊ/. Compare this with the three-vowel-quality system of Tunis Arabic, which also lacks diphthongs ([Gibson 2011](#)).

Since the Italo-Romance languages have vowel systems of a similar richness to Maltese, one might assume that this proliferation of vowel phonemes is a straightforward case of transfer. This is, in general, not the case, however. The majority of new phonemic distinctions are at least partially the result of the loss of emphatic consonants and of *ʔ,⁶ which led to the phonemicization of vowel qualities that were previously merely allophonic. Note also that the innovative lax close front long vowel /ɪ:/ is apparently an entirely internal development – the outcome of an extreme raising of the front allophone of *ā (so-called *imāla*), as in *ktieb* /ktr:b/ ‘book’ < *kitāb*.

Following Krier (1976: 21–22), we can nevertheless point to three innovations in this domain which do seem to be the direct result of lexical borrowing from Italo-Romance.

Krier (1976: 21) points out first of all that, of the five short vowels, only four of them /ɪ, ɛ, ɐ, ɔ/ appear in all positions in Arabic-derived lexicon. In this portion of the lexicon, /ʊ/ occurs only in final position in unstressed syllables, with the single exception of *kull* ‘all’. Were it not for the (extensive) Italo-Romance component of the Maltese lexicon, therefore, we can say that the distinction between [ɔ] and [ʊ] would remain allophonic, as it is in Tunis Arabic. As it is, the two sounds should probably be considered phonemically distinct in Maltese. Although minimal pairs are hard to find, possible examples include *punt* ‘point’ vs. *pont* ‘bridge’ and *lotto* ‘lottery’ vs. *luttu* ‘mourning’.⁷

⁶These latter changes are themselves, however, arguably contact-induced – see §3.1.1.2.

⁷Our thanks to Michael Spagnol for suggesting these examples.

Among the long vowels, the presence of /ɛ:/ and /ɔ:/ phonemes in Maltese is also largely attributable to Italo-Romance loans containing these sounds. Although /ē/ and /ō/ occur in certain Tunisian Arabic varieties (Gibson 2011; Herin & Zammit 2017), these are the result of historical monophthongization of the original *ay and *aw diphthongs. The Maltese reflexes of these sounds remain diphthongs, as in *sejġ* /sɛɪf/ ‘sword’ and *lewn* /lɛʊn/ ‘colour’. Other than in cases of compensatory lengthening in items where the consonants represented by <gh> and <h> have been lost (see §3.1.1.2), /ɛ:/ and /ɔ:/ only occur in the non-Arabic component of the Maltese lexicon, as in *żero* /'zɛ:rɔ/ ‘zero’ and *froġa* /'frɔ:dʒɐ/ ‘omelette’.

To these three contact-induced monophthongal innovations we can add one new contact-induced diphthong: /ɔɪ/. Mifsud (2011) points out that this occurs only in non-Arabic lexical items (e.g. *vojt* /vɔɪt/ ‘empty space’) in standard Maltese.

In summary, then, the majority of innovative vowel phonemes in Maltese are not the direct result of transfer, but the three new monophthongal phonemes whose emergence is (at least partially) contact-induced, combine to create a nearly symmetrical system in which all five short vowel phonemes have a long counterpart.

3.1.3 Intonation

Despite pioneering work by Alexandra Vella (e.g. Vella 1994; 2003; 2009; Grice et al. 2019), the study of intonation in Maltese, as in most non-Indo-European languages, remains in its infancy (cf. Hellmuth, this volume). Impressionistically speaking, the tunes that can be heard in Maltese (and Maltese English) speech are highly distinctive, and often quite unlike those of the Mediterranean Arabic dialects. Several studies have demonstrated that intonation patterns are highly susceptible to transfer in language contact situations, especially through imposition by source-language-dominant speakers (see the studies of Spanish intonation by O'Rourke 2005; Gabriel & Kireva 2014). Interestingly, however, this appears to be less true for the tunes associated with polar interrogatives, at least in the varieties of Spanish described by the aforementioned authors, presumably because of the importance of intonation in establishing interrogative force in the absence of syntactic cues in this language. What data we have on this issue for Maltese fits rather neatly into this larger picture. According to Vella (2003), the intonational patterns of Maltese late-focus declaratives on the one hand, and wh-interrogatives on the other, pattern with Palermo Sicilian and Tuscan Italian respectively, while that of Maltese polar interrogatives more closely resembles counterparts in Arabic dialects.

It seems safe to assume that imposition by native speakers of Italo-Romance varieties is the primary cause of the similarities between Maltese and Italo-Romance intonation, but borrowing by Maltese-dominant bilinguals should not be ruled out as an additional factor.

3.2 Morphology

3.2.1 Nouns and adjectives

3.2.1.1 Inflection

It has been shown (e.g. [Gardani 2012](#); [Seifart 2017](#)) that plural affixes are, with case affixes, the most widely transferred inflectional morphemes. Maltese conforms neatly to the general crosslinguistic picture: it has acquired plural morphemes from Sicilian and English and little in the way of other inflectional morphology (but see §3.2.2).⁸

In addition to a rich array of stem-altering (so-called “broken”) plural patterns, most of which also serve as the plurals of at least some items of Italo-Romance or, more rarely, English origin (see [Spagnol 2011](#) for details), Maltese has six plural suffixes: *-in*, *-a*, *-iet*, *-ijiet*, *-i*, and *-s*.⁹ Of these, *-in*, and *-iet* are straightforward retentions from Arabic (nevertheless extended to numerous non-Arabic items), *-i* and *-s* are straightforward cases of indirect affix borrowing (in the sense of [Seifart 2015](#)), and *-a*, and *-ijiet* arguably involve a subtle interplay of internal and externally-caused developments.

The most recently borrowed plural suffix is the English-derived *-s*. This occurs exclusively with bases borrowed from English, and may be considered only partially integrated into monolingual Maltese (to the extent that such a thing exists; see §2.2), in that it often alternates optionally with *-ijiet* in items such as *kejk* ‘cake’ (pl. *kejkijiet* ~ *kejks*). There are, however, a number of reasonably frequent items (e.g. *friżer* ‘freezer’) which appear never to take a plural suffix other than *-s*.

The Sicilian-derived suffix *-i* can mark the plural of a far higher proportion of Maltese nouns than can *-s*, and is demonstrably better integrated into the Maltese inflectional system. In addition to marking the plural of Sicilian-derived nouns which also take *-i* (e.g. *xkupa* ‘broom’ < Sicilian *scupa* (pl. *scupi*); *fjakk* ‘weak’ <

⁸One should note also, however, the appearance in a couple of items of a singulative suffix *-u*, apparently borrowed from Sicilian. Borg (1994: 57) cites *wiżż-u* ‘geese-SING’, *dud-u* ‘worms-SING’, and *ful-u* ‘beans-SING’.

⁹There are also one or two examples of zero plurals, e.g. *martri* ‘martyr(s)’.

Sicilian *fiaccu* (pl. *fiacchi*)), it has also been extended to: Italian-derived nouns (including those with a plural in *-e* in Italian, e.g. *statwa* ‘statue’ < Italian *statua* (pl. *statue*)); nouns from other Romance languages (e.g. *pitrava* ‘beetroot’ < French *betterave* with Ø-plural (orthographic *-s*)); English-derived nouns (e.g. *jard* ‘yard (unit of distance)’), and even a few Arabic-derived nouns (e.g. *saff* ‘layer’ < *ṣaff* ‘row’, *samm* ‘very hard’ < *ʔaṣamm* ‘deaf, hard’).

Arabic and Sicilian coincidentally have an identical less frequently used plural (or collective) suffix *-a*, as in Arabic *mārra* ‘passers-by’ (singular *mārr*) and Sicilian *libbra* ‘books’ (singular *libbru*). A plural suffix of this form also occurs in Maltese, with nouns of both Arabic and Italo-Romance origin (e.g. *kittieba* ‘writers’ < Arabic *kattāb*; *nutara* ‘notaries’ < Italian *notaro*). Evidence that this is perceived and treated as a single morpheme rather than two homophonous items comes from the fact that the restriction of this suffix to groups of people in Arabic applies also to the Italo-Romance part of the Maltese lexicon (Mifsud 2011).

A curious feature of Maltese plural morphology from a comparative Arabic perspective is the very frequent suffix *-ijiet* (*-jiet* after certain vowel-final stems), as in *postijiet* ‘places’ (singular *post*) and *ommijiet* ‘mothers’ (singular *omm*). While clearly based on the Arabic-derived suffix *-iet* (< Arabic *-āt*, with characteristic Maltese *imāla*), the provenance of the initial *-ij-* is not obvious. Mifsud (2011) plausibly suggests that *-ijiet* as a whole is “derived from the plural of verbal nouns with a weak final radical, like *tigrijiet* ‘races’, *tiswijiet* ‘repairs’”, but Geary (2017) makes a strong case that the large influx into Maltese of Italo-Romance nouns whose singulars ended in *-i* (e.g. *affari* ‘affair, matter’ < Sicilian *affari* or Italian *affare*) was instrumental in the emergence of this morpheme. On this account Maltese speakers originally pluralized such words with *-iet*, with glide-insertion an automatic phonological consequence of the juncture of a vowel-final stem and a vowel-initial suffix. Later, according to Geary, the whole string *-ijiet* was reanalysed as constituting the marker of plurality, and this new plural suffix was extended to consonant-final stems, including Arabic-derived items of basic vocabulary such as *omm* ‘mother’ and *art* ‘land’.¹⁰

¹⁰Geary’s contact-induced scenario for the emergence of this suffix may not be the whole story, however. Evidence on this point comes from Arabic loanwords in Siwa Berber. Souag (2013: 74) lists a number of examples of Arabic-origin nouns whose plural is formed by adding a suffix *-iyyat* (e.g. *ṣhilfa* ‘turtle’, pl. *ṣhilfiyyat*), despite the fact that both Classical Arabic and present-day Egyptian Arabic lack plurals of this type. Siwa Berber must therefore have borrowed these items and their pluralization strategy from some early form of (eastern) Maghrebi Arabic, suggesting that the presence in Maltese of the *-ijiet* suffix is, at least to some extent, an Arabic-internal development that predates the large-scale borrowing of Italo-Romance nouns into Maltese.

3.2.1.2 Derivation

Maltese displays a rich array of derivational suffixes borrowed (presumably initially as part of polymorphemic lexical items) from Italo-Romance. A definitive list of these has not been provided to date, but [Saade \(2019\)](#) offers a detailed typology of such items, of which we present a simplified version here, drawing also on examples from [Brincat & Mifsud \(2015\)](#), and focusing just on the nominal, adjectival and adverbial domains (see §3.2.2.2 for borrowed participial morphology).

First of all, there are at least twenty suffixes, such as the nominalizer *-zzjoni*, which, though relatively frequent, only occur in items clearly borrowed wholesale from Italo-Romance (e.g. *dikjarazzjoni* ‘declaration’ < Italian *dichiarazione*) or in coinages which, in a process that is relatively common in Maltese, represent borrowings from English that are adapted to fit the phonology and morphology of Romance-influenced Maltese, as in *esplojtazzjoni* ‘exploitation’ (cf. [Gatt & Fabri 2018](#)). Given this restriction, there must be some doubt as to whether one can regard the suffixes themselves as borrowed, or only the polymorphemic items in which they occur.

Secondly, there are a number of borrowed suffixes which are sufficiently well integrated that they can attach to Arabic-derived bases. Examples include:

- ata*, e.g. *xemxata* ‘sunstroke’ (*xemx* ‘sun’)
- ezza*, e.g. *mqarebezza* ‘naughtiness’ (*mqareb* ‘naughty’)
- un* (< Sicilian *-uni*, Italian *-one*), e.g. *ħmarun* ‘great fool’ (*ħmar* ‘donkey’)

Finally, there is at least one borrowed suffix *-tura*, which forms a single-instance verbal noun, whose integration can be seen from the fact that it attaches to productively to English bases, as in *ċekkjatura* ‘an instance of checking’ or *weldjatura* ‘an instance of welding’.

3.2.2 Verbs

3.2.2.1 Loaned verbs

Maltese has borrowed a large number of verbs from Sicilian and Italian, and more recently a smaller number from English. The chief interest in these borrowings lies in the way in which they have been integrated into the Maltese inflectional and derivational verbal paradigms. An in-depth study of this phenomenon was provided by [Mifsud \(1995\)](#), who distinguished the following four types of loaned verbs:

Type A: Full integration into Semitic Maltese sound verbs

Christopher Lucas & Slavomír Čéplö

Type B: Full integration into Semitic Maltese weak-final verbs

Type C: Undigested Romance stems with a weak-final conjugation

Type D: Undigested English stems

Mifsud (1995: 58) points out that most (perhaps all) Type A verbs are so-called “second generation” loans, whereby a nominal or adjectival form has been borrowed, a root extracted from it, and a verb formed on this root, as in *pitter* ‘to paint’ – a denominal derivation from *pittur* ‘painter’, borrowed from Sicilian *pitturi* (and supported by Italian *pittore*). Such items do not, therefore, represent genuine cases of transfer of verbs, and are reminiscent of similar coinages in other Arabic varieties (e.g. *fabrak* ‘to fabricate’). In Arabic as in Maltese, such items are overwhelmingly restricted to the denominal verbal stems II and V of trilateral roots and I and II of quadrilateral roots (CVCCVC and tCVCCVC).

In contrast to Type A, Mifsud’s Types B and C are genuine cases of loaned verbs. Mifsud (1995: 110–116) shows that the imperative (rather than the homophonous 3SG present, or any other verb forms) was the most likely base form of the Romance models on which the Maltese loaned verbs were created.¹¹ In both Italian and Sicilian all verbs in the imperative end in either *-i* or *-a*. As it happens, Maltese weak-final verbs (in which the final radical element is a vowel rather than a consonant) also all end in either /i/ or /e/ in the imperfect and imperative singular, depending on which of the two weak-final conjugation classes they fall into. This coincidence resulted in borrowed Romance verbs being integrated into one of the these two weak-final classes, as in *kanta* ‘he sang’, *jkanta* ‘he sings’ (< Sicilian/Italian imperative *canta*); and *serva* ‘he served’, *jservi* ‘he serves’ (< Sicilian/Italian imperative *servi*).

The difference between verbs of Types B and C is that the former are analysed as having root-and-pattern morphology, with a trilateral or quadrilateral root, whereas Type C are borrowed as a concatenative stem without a root. This can be seen from the fact that Type B verbs can give rise to new verbs with the same root in other verbal stems, as in *kompla* ‘to continue’, *tkompla* ‘to be continued’ (< Sicilian *cumpliri* ‘to finish’), whereas Type C verbs cannot.

Another difference between Types B and C is that no Type C verb begins with a single (ungeminated) consonant, whereas most Type B verbs do. In fact, apart from certain well-defined exceptions (see Mifsud 1995: 152), all Type C verbs begin with a geminate consonant, as in *ffolla* ‘to crowd’ < Italian *affollare*. What exactly was the combination of historical factors that gave rise to this synchronic

¹¹This parallels the situation in Arabic-based pidgins and creoles, for which Versteegh (2014) shows that verbs generally appear to derive from infinitives in the lexifier varieties.

state of affairs is a complex matter (see Mifsud 1995: 158–168 for discussion), but the key point to note is that at least some of the instances of initial gemination in Type C verbs are apparently not attributable to phonological properties of the source item (e.g. *pprova* ‘to try’ < Italian *provare*). It seems that speakers of Maltese came to feel that all loan verbs must have an initial geminate consonant, whether or not this was actually true of the item being borrowed.

This state of affairs manifests itself rather spectacularly in more recent borrowings from English (Type D verbs), in which initial consonants are duly geminated (despite this never being the case in the English source items), but which also fall into the conjugation class of weak-final verbs, as in *ddawnlowdja* ‘to download’. What underlies this treatment of loans from English seems to be a type of reanalysis, which we can sketch as follows. In the initial stage, verbs without roots (not necessarily identifiable to speakers as loans from Italo-Romance) are analysed as falling into the weak-final conjugation class because they have a stem-final vowel. But since all verbs without roots (at this pre-English stage) have a stem-final vowel, it is possible to view the lack of a root, not the presence of a stem-final vowel, as the reason that loan verbs obligatorily fall into the weak-final conjugation class; and it seems that speakers indeed made this reanalysis. In a parallel development, initial consonant gemination also came to be seen an obligatory feature of the class of verbs lacking a root. As a result of these developments, when a verb is borrowed from English, because it lacks a root its initial consonant is geminated and it is conjugated as a weak-final verb, regardless of whether it has a stem-final vowel.¹²

3.2.2.2 Participles

Unsurprisingly, one of the additional ways in which Type A verbs differ from the remaining three classes of loaned verbs is the formation of passive participles: in Type A verbs, passive participles are formed in accordance with the Semitic pattern for the respective derived stem, e.g. *pejjeġ* ‘to smoke’ (stem II, from Italian *pipa* ‘pipe’) produces *mpejjeġ* ‘smoked’ (Mifsud 1995: 70). In contrast, some Type B verbs allow for the formation of a passive participle using Romance suffixes (Mifsud 1995: 127–133), and this is the sole option for Type C and even Type D verbs: for Type C verbs, the choice of the actual suffix depends on the origi-

¹²In addition, virtually all Type D verbs insert a palatal glide between the borrowed stem and the added weak-final vowel, as in *pparkja* ‘to park’. Similarly to the initial gemination and weak-final inflection of Type D verbs, this glide insertion must be the result of analogical extension from numerous glide-final borrowed Romance verbs, e.g. *rdoppja* ‘to double’ < Italian *raddoppiare*. See Mifsud (1995: 225–236) for a detailed discussion.

Christopher Lucas & Slavomír Čéplö

nal form of the verb and, in some cases, the path of borrowing (see below). For Type D verbs borrowed from English, the suffix *-at* is the only productive way to form a passive participle (e.g. *inxurjat* ‘insured’) with *spellut* ‘spelled’ as the only exception (Mifsud 1995: 248).

And finally, there are two distinct classes of Type B and C verbs which can each derive two passive participles. In the first class, one participle is derived from the weak (regular) form root and the other derived from the strong one, e.g. *konfondut* ‘confused’ vs. *konfuż* (Mifsud 1995: 134). In the second class, one participle is derived using the Sicilian suffix *-ut*, the other using the Italian-derived suffix *-it*, e.g. *preferut* ‘preferred’ vs. *preferit* (Mifsud 1995: 230). The reason for these doublets is largely sociolinguistic: the variability of the first class echoes a similar situation in Italian dialects (Mifsud 1995: 134); that of the second class reflects a situation whereby the loaned verb effectively has two sources, spoken Sicilian and standard (Tuscan) Italian.

3.3 Syntax

3.3.1 Phrase syntax

3.3.1.1 Word order

The expansion of Maltese lexicon with items borrowed from Sicilian and Italian had a profound effect on the syntax of Maltese. The primary example of this is word order within the noun phrase, involving the order of adjectives and their heads. In Arabic, adjectives (with the exception of comparatives, superlatives and a number of specific cases) follow their heads. This is largely true of Italian adjectives as well, with the exception of a small subclass some grammars term “specificational adjectives” (e.g. Maiden & Robustelli 2007: 55–56), such as *stesso* ‘same’ and *certo* ‘certain’, which precede their head. Such adjectives borrowed into Maltese retained their syntactic properties, as with the pre-nominal *ċertu* (< Sic. *certu*) in (3).

- (3) [BCv3: it-torca.8685]
 Kien bniedem ta’ ċerta personalità.
 be.PR.F.3SG.M person GEN certain.F personality
 ‘He was a person with a certain personality.’

In Italian, specificational adjectives to a large extent overlap with a class of adjectives that perform double duty as quantifiers (or perhaps determiners) and vary their position according to their respective roles: Adj–N for quantifiers, N–Adj for adjectives. One could argue that it is in the former function that they were

borrowed into Maltese and thus should be considered quantifiers or determiners rather than adjectives, especially in light of the fact that they are (for the most part) in complementary distribution with the definite article, as determiners and quantifiers are. Determiners and quantifiers in Maltese precede their heads (as with the definite article *il-*, *kull* ‘all’, *xi* ‘some’ etc.).

There are three arguments against such an account: first of all, borrowed pre-nominal specificational adjectives actually fall into two classes, where members of the first, such as *ċertu* ‘certain’, *diversi* ‘diverse’ (< It. *diverso*) or *varju* ‘various’ (< Sic. *varju*), do not (for the most part) allow the definite article. In contrast, words in the second class such as *stess* ‘same’ (< It. *stesso*) or *uniku* ‘unique’ (< Sic. *uniku*) predominantly co-occur with the definite article when pre-nominal. The same, incidentally, is true of the etymologically Arabic pre-nominal quantifier *ebda* ‘no, none’.

Secondly, there are morphological considerations: pre-nominal specificational adjectives of both types mark gender and/or number (*varju* for the first, *uniku* for the second) like Maltese adjectives do; Maltese determiners and quantifiers do not inflect for either gender or number.¹³

The final argument against considering borrowed pre-nominal specificational adjectives as being borrowed into the slot for determiners involves ordinal numerals. In Italian, these also fall into the subclass of prenominal specificational adjectives (Maiden & Robustelli 2007: 55) and thus precede their head. The same is invariably true of Maltese ordinal numerals, as with *ewwel* in (4).

- (4) [BCv3: l-orizzont.64586]
 Wara l-ewwel sena
 after DEF-first year
 ‘After the first year’

In North African Arabic, ordinal numerals can either precede or follow their heads, but when they precede them, they never take the definite article, even when the noun phrase is semantically definite (see e.g. Ritt-Benmimoun 2014: 284 for Tunisian Arabic). In contrast, Maltese never allows its ordinal numerals to follow their heads, and the definite article is obligatory.

All these arguments, including the comparison with related Arabic varieties, suggest that the pre-nominal position of some adjectives and ordinal numerals in Maltese is due to transfer under recipient-language agentivity from Italian.

¹³With the exception of the very specific category of demonstrative pronouns where gender and number are marked not by affixes, but rather a form of suppletion.

3.3.1.2 The analytical passive

As with adjectives (§3.3.1.1), lexical borrowings from Italo-Romance have also had a significant impact on the syntax of Maltese verbs. One of the most conspicuous consequences of this development involves the passive voice: as Romance-origin verbs cannot generally form one of the passive derived verbal stems (but see §3.2.2.1), they brought with them their Romance syntax and thus a new type of passive construction arose in Maltese – the analytical passive.

In Maltese, there are two types of analytical passive construction containing a passive participle: the so-called “dynamic passive” (Vanhove 1993: 321–324; Borg & Azzopardi-Alexander 1997: 214), which combines passive participles with the passive auxiliary *gie* ‘to come’; and the so-called “stative passive” (Borg & Azzopardi-Alexander 1997: 214, Vanhove 1993: 318–320), which has the same structure as copular clauses (see §3.3.2.3), the only difference being that stative passive constructions can feature an agentive NP introduced by the preposition *minn* ‘from’ (see Čéplö 2018: 104–107 for a detailed analysis).

The stative passive can be viewed as an extension of the structurally identical construction which is sporadically attested already in Classical Arabic (Ullmann 1989: 76–84), but becomes quite prominent in Christian Arabic documents at least as early as tenth century, where, incidentally, it gained prominence under influence from Aramaic and Greek (Blau 1967: 424).

The dynamic passive (5), on the other hand, is a straightforward calque on either Italian or Sicilian, where a construction featuring a verb semantically equivalent to *gie* ‘to come’ – *venire* in Italian – combines with a past participle (cf. Manfredi, this volume).

- (5) [MUDTv1: 30_01P05]
 Kif diġà ghedt, gie pprezentat il-kuntratt.
 as already say.PRF.1SG come.PRF.3SG.M present.PTCP.PASS.M DEF-contract
 ‘As I already said, the contract was presented.’

While the dynamic passive must have originally functioned to fill a hole in the verbal system of Maltese by providing a way to passivize Romance verbs, it has meanwhile spread to include native verbs as well, as with *ta* ‘to give’ (< $\sqrt{\text{fty}}$) in (6).

- (6) [BCv3: inewsmalta-ott.29.2013.1257-11045]
 It-tagħrif gie mogħti mill-Ministru
 DEF-information come.PRF.3SG.M give.PTCP.PASS from.DEF-minister

Konrad Mizzi.

Konrad Mizzi.

‘The information was given by Minister Konrad Mizzi.’

3.3.1.3 Modality

Another clear-cut example of grammatical calquing comes from the domain of modality and involves the pseudoverb *għand-*. In Maltese, its primary function is that of a possessive (7), as is the case with its cognates *sind-/sand-* in many Arabic varieties.

- (7) [MUDTv1: 22_02J03]
 M’ għandi xejn kontri-hom.
 NEG have.1SG nothing against-3PL
 ‘I have nothing against them.’

In addition to this, however, the Maltese *għand-* has also taken on a function as a deontic modal of weak obligation ‘should, ought’ taking verbal complement, as in (8).¹⁴

- (8) [MUDTv1: 22_02J03]
 Naqbel li għandhom jivvutaw aktar nies.
 agree.IMPF.1SG COMP have.3PL vote.IMPF.3PL more people.
 ‘I agree that more people should vote.’

The use of *għand-* in this kind of modal function appears to be unique to Maltese; not even Cypriot Maronite Arabic with its many parallels to Maltese (on which see below) exhibits the same behavior for its cognate *sint-* (Borg 2004: 346) and uses a different verb, *salah/pkyislah* (Borg 2004: 323), as the default deontic modal. The Maltese development must therefore be another calque, since the basic possessive verb of Sicilian, *aviri*, also doubles as a deontic modal, as in (9).

- (9) Sicilian (Piccitto 1977: 340)
 Cci l’ài a-ddiri a-tto patri.
 DAT.3SG.M OBJ.3SG.M-have.PRES.1SG to-say.INF DAT-2SG.M father
 ‘I have to say it to your father.’

¹⁴ *għand-* is the only Maltese pseudoverb (and verb) which exhibits a three-way distinction between present (*għand-*), past (*kell-*) and future/habitual (*ikoll-*) forms; all can occur in the modal function.

3.3.2 Sentence syntax

3.3.2.1 Differential object marking

Differential object marking (DOM) is a phenomenon where direct objects are marked according to some combination of semantic and pragmatic properties of the object in question. In Spanish, for example, objects denoting humans (and equivalent entities) are marked by the particle *a*, originally a directional preposition. DOM is a phenomenon attested cross-linguistically (see [Khan 1984](#) for Semitic languages), including in varieties of Arabic like Levantine and Iraqi Arabic ([Coghill 2014](#) and references therein) and in Andalusian Arabic (*A descriptive and comparative grammar of Andalusian Arabic* 2013: 108).

DOM is a well-documented feature of Maltese morphosyntax and largely conforms to the Spanish prototype: in general, both pronominal and nominal direct objects denoting entities high in the “animacy hierarchy” ([Borg & Azzopardi-Alexander 1997](#): 55) take the object marker *lil* (10), which also does double duty as the indirect object marker for all objects. Inanimate objects do not take *lil* (11).

- (10) [BCv3: ilgensillum.2011-Mejju-22.8230]
 Min jara lili jara lil Missier-i
 who see.IMPF.3SG.M OBJ.1sg see.IMPF.3SG.M OBJ father-OBL.1SG
 ‘Who looks at me, looks at my Father’

- (11) [BCv3: l-emigrant]
 Min jara orrizzonti godda u min baħħ.
 who see.IMPF.3SG.M horizon.PL new.PL and who void.
 ‘Some see new horizons, some see a void.’

[Döhla \(2016\)](#) examines DOM in Maltese in some detail and arrives at the conclusion that while there is “a certain predisposition for object marking in general within pan-Arabic grammar” ([Döhla 2016](#): 169), Maltese DOM cannot be ascribed to purely internal developments within Arabic. A striking feature of the Arabic varieties that exhibit DOM is that they were all in prolonged contact with other languages: Aramaic for Levantine and Iraqi Arabic (and, by extension, for Cypriot Maronite Arabic, cf. [Borg 2004](#): 412), Romance for Andalusian Arabic and Maltese. In the case of Maltese, said Romance variety is Sicilian, where the object marker *a* performs the same double duty as the Maltese *lil*, and DOM in both languages shows a number of remarkable similarities: in both Sicilian and Maltese, DOM is primarily triggered “by humanness along with definiteness/referentiality” ([Iemmolo 2010](#): 257, in reference to Sicilian), it is obligatory with personal pronouns,

but optional with plural “kinship terms and human common nouns” and disallowed with “(in)animate and indefinite non-specific nouns” (Iemmolo 2010: 257, again in reference to Sicilian), as exemplified by the non-specific Maltese *nies* ‘people’ in (12).

- (12) [BCv3: l-orizzont.41390]
 Min irid jara nies jgħixu hekk?
 who want.IMPF.3SG.M see.IMPF.3SG.M people live.IMPF.3PL.M thus
 ‘Who wants to see people live like that?’

In Maltese DOM, then, we have an instance of what Manfredi (this volume) labels “calquing of polyfunctionality of grammatical items inducing syntactic change”: Maltese acquired a rule of DOM as a result of the indirect object marker *lil* inheriting the dual function of its Sicilian equivalent *a*. It is clear that this is a contact-induced change. But since with this and the similar changes discussed below there is no transfer of lexical matter, it seems impossible at present to judge whether they are the result of borrowing or imposition, or whether they were actuated by speakers for whom neither the RL or the SL were dominant, in the process that Lucas (2015) calls “convergence”.

3.3.2.2 Clitic doubling (proper)

The existence of various reduplicative phenomena associated with direct and indirect clitic pronouns in Maltese has been noted at least since Sutcliffe (1936: 179), who identifies what classical tradition refers to as *nominativus pendens*. This analysis has been elaborated on by Fabri (1993), Borg & Azzopardi-Alexander (1997) and Fabri & Borg (2002), primarily in the context of pragmatically determined constituent order variation, especially topicalization. Building on these works and the analysis of Maltese clitics by Camilleri (2011), Čéplö (2014) notes that in addition to these phenomena, which in one way or another entail dislocation, there exists in Maltese another related phenomenon, where lexical objects and clitic pronouns co-occur, but without the dislocation of the lexical object. This phenomenon, termed Clitic Doubling Proper to distinguish it from similar constructions (see Krapova & Cinque 2008 for a detailed analysis), involves the co-occurrence of a lexical object and the clitic with the object in situ, which in Maltese is after the verb (see Čéplö 2018). Maltese Clitic Doubling Proper occurs with both direct (13) and indirect objects (14).

- (13) [BCv3: l-orizzont.36758]

Christopher Lucas & Slavomír Čéplö

Ftit nies jafu-ha l-istorja marbuta ma'
 few people know.IMPV.3PL.M-3SG.F DEF-history connected.SG.F with
 dan il-proġett tant sabiħ.
 DEM.SG.M DEF-project such beautiful
 'Few people know the history connected with such a beautiful project.'

- (14) [BCv3: 20020313_714d_par]
 Hekk qed ngħidu-lhom lil dawn in-nies f' pajjiż-na.
 thus PROG say.IMPV.1PL-DAT.3PL DAT DEM.PL DEF-people in country-1PL
 'This is what we say to these people in our country.'

Unlike various types of dislocation with resumptive clitic pronouns which are quite common in European languages (see e.g. [de Cat 2010](#)), Clitic Doubling Proper is a much rarer phenomenon; in Europe, it is largely confined to the Balkan *Sprachbund* ([Friedman 2008](#)) and some Romance languages outside of the Balkans, like Spanish ([Zagona 2002: 7](#)) and varieties of Italian ([Russi 2008: 231–233](#)). The phenomenon is also attested in Semitic languages ([Khan 1984](#)), including Arabic, where it was studied in detail by [Souag \(2017\)](#). Comparing Clitic Doubling Proper in various varieties of Arabic including Maltese, [Souag \(2017: 57\)](#) notes parallels between Maltese and some varieties of Algerian Arabic, especially in regard to the doubling of indirect objects. Ultimately, however, he arrives at the conclusion that Maltese Clitic Doubling Proper “has little in common with any other Arabic variety examined, but closely resembles that found in Sicilian” ([Souag 2017: 60](#)). This suggests that here too we have a contact-induced change, this time of the sort that Manfredi (this volume) labels “narrow syntactic calquing”, that is, without any accompanying calque of lexical items.

3.3.2.3 Copular constructions

In Maltese, there are four types of copular clauses ([Borg & Azzopardi-Alexander 1997: 53](#)):¹⁵

- Type 1: No copula
- Type 2: The verb *kien* as the copula
- Type 3: Personal pronoun as the copula

¹⁵In addition to these, [Borg \(1987–88\)](#) and [Borg & Spagnol \(2015\)](#) also describe the copular function of the verb *jinsab* ‘to be found’. This being a finite verb, both [Borg & Azzopardi-Alexander \(1997: 53\)](#) and [Čéplö \(2018: 99–104\)](#) exclude this type of clause, as well as similar ones, such as those featuring the verb *sar* ‘to become’, from the category of copular clauses.

Type 4: Present participle *qiegħed* as the copula

Type 1 describes what traditional grammars of Semitic languages refer to as nominal sentences; Type 2 then contains their counterparts with a past reference. Types 3 and 4, while not without parallel in other varieties Arabic,¹⁶ feature much more prominently in Maltese. This is especially true of Type 3 copular clauses, which involve the use of a personal pronoun as the copula (15).

- (15) [BCv3: 2010 Immanuel Mifsud - Fl-Isem tal-Missier (U tal-Iben)]

Din hi omm-ok.

this.F 3SG.F mother-2SG

‘This is your mother.’

Similar copular constructions to that illustrated in (15) have been described for several Maghrebi varieties (cf. [Vanhove 1993](#): 355), but Maltese stands apart in terms of the frequency with which Type 3 constructions occur: in MUDTV1, for example, 110 non-negative copular clauses are of Type 1; 181 are Type 3. In this, Maltese Type 3 copular clauses are comparable to equivalent copular constructions in Anatolian Arabic (see [Lahdo 2009](#): 172–173 for Tillo Arabic and the references therein, as well as Akkuş, this volume), Andalusian Arabic (*A descriptive and comparative grammar of Andalusī Arabic* 2013: 105), and especially Cypriot Maronite Arabic ([Borg 1985](#): 135; Walter, this volume), where they are but one piece of evidence linking Cypriot Maronite Arabic to *qaltu* dialects ([Borg 2004](#): 31). The conclusion to be drawn here is the same as for DOM and Clitic Doubling Proper above: it is no coincidence that these copular constructions are in wide use and the copular construction of choice especially in varieties of Arabic which have been under contact influence from languages with a mandatory copula – Turkish for Anatolian Arabic, Spanish for Andalusian Arabic, Greek for Cypriot Maronite Arabic, and Italian for Maltese. Whether the origin of such constructions can be traced to a feature in (one of) these dialects’ Old Arabic ancestors, or whether they came about through parallel development, contact undoubtedly triggered the widespread adoption of such constructions in these varieties of Arabic.

3.4 Lexicon

3.4.1 Major sources

That Maltese contains large numbers of loanwords from Romance and English is a fact immediately obvious to even the most casual observer. Over the years,

¹⁶See the analysis of Type 4 copulas in [Camilleri & Sadler \(2018\)](#).

Christopher Lucas & Slavomír Čéplö

there have been a number of attempts to quantify the influence of other languages on Maltese by providing a classification of lexemes by their origin. The earliest, Fenech (1978: 216–217), compiled such statistics for journalistic Maltese, but also provided a comparison to literary and spoken Maltese (albeit using a very small data sample). Brincat analyzed the etymological composition of entries in Aquilina’s dictionary, first examining the origin of 34,968 out of all 39,149 headwords (Brincat 1996: 115) and then applying the same analysis to the entire list (Brincat 2011: 407); Mifsud & Borg (1997) did the same with the vocabulary contained in an introductory textbook of Maltese as a foreign language. In 2006, Bovington & Dalli (2006) analyzed the etymology of lexical items in a 1000-word sample obtained from a corpus of Maltese and, most recently, Comrie & Spagnol (2016: 318) did the same on a list of 1500 “lexical meanings” within the framework of the *Loanwords in the world’s languages* project (Haspelmath & Tadmor 2009). Figure 1 summarizes all these findings.

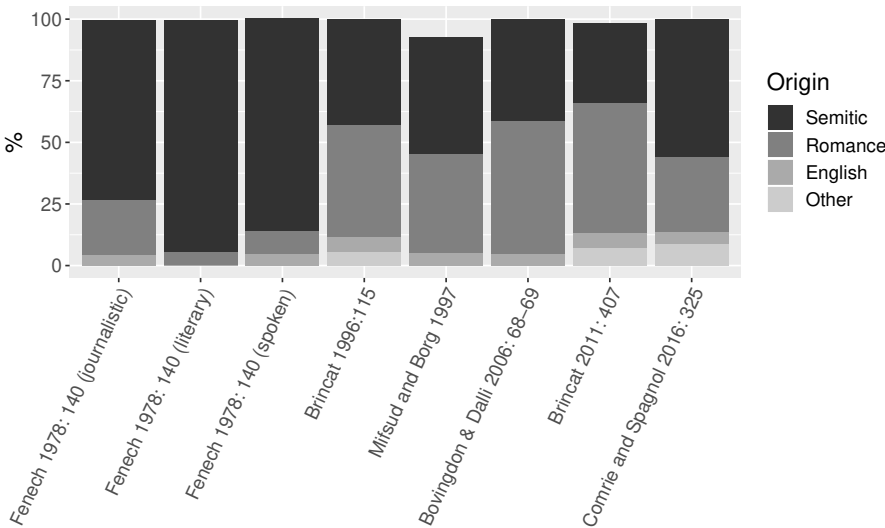


Figure 1: A summary of previous studies of the composition of the Maltese lexicon

The primary explanation for the sharp differences between these analyses is methodology: while Fenech (1978) analyzes entire texts and thus counts tokens, Brincat (1996) (including its updated version in Brincat 2011) and Bovington & Dalli (2006) analyze lists of unique words, i.e. types. The later is also true of Mifsud & Borg (1997) and Comrie & Spagnol (2016), except where Brincat (1996) uses

dictionary data and Bovingdon & Dalli (2006) corpus data, Mifsud & Borg (1997) employ a list of lexical items with high frequency of use in daily communication and Comrie & Spagnol (2016) base their analysis on a list compiled for the purposes of cross-linguistic comparison. The high ratio of words of Semitic origin in token-based analyses is thus due to the prevalence of function words, which are overwhelmingly Arabic. The type-based analyses then provide a somewhat more accurate picture of the lexicon as a whole, even though they are not without their problems. Chief among these is the issue of what exactly counts as type, especially with regard to productive derivational affixes, e.g. whether all the words with the prefix *anti-* count as distinct types or not.

In addition to general analyses, both Bovingdon & Dalli (2006) and Comrie & Spagnol (2016) also provide breakdowns for individual parts of speech. Unfortunately, these analyses are not comparable, as each has a different focus: Bovingdon & Dali (2006: 71) are interested in the composition of each etymological stock by word class (Table 3).

Table 3: Source language component of Maltese by word class (Bovingdon & Dalli 2006: 71).

Origin	Function words	Verbs	Adjectives	Nouns	Adverbs	Pronouns
Semitic	3%	70%	2%	21%	2%	2%
Romance	0%	38%	11%	48%	3%	0%
English	0%	29%	8%	63%	0%	0%

In contrast, Comrie & Spagnol (2016: 328) focus on the composition of individual word classes by their origin (Table 4).¹⁷

¹⁷The details of Comrie & Spagnol’s (2016) methodology mean that loans in their dataset come from Romance and English but not from any other languages. The category we label “Misc.” in Tables 4 and 5 encompasses those meanings in the *Loanwords in the world’s languages* 1500-item set which have no corresponding single-word Maltese lexical item, and those where the etymology is at present unknown, or where the item in question is an innovative Maltese-internal coinage.

Table 4: Word class composition by source language (Comrie & Spagnol 2016: 328)

Word class	Arabic	Romance	English	Misc.
Function words	84.7%	6.2%	0%	9.1%
Verbs	75.3%	14.1%	1.3%	9.2%
Adjectives	65.2%	28.5%	0.3%	6.0%
Nouns	44.7%	39.6%	7.2%	8.6%

Comrie & Spagnol (2016) also provide a breakdown of their data by semantic field, permitting a comparison of the domains in which Romance versus English loans are more or less prominent. A number of generalizations can be made here (see Table 5 for a summary), though ultimately they all follow naturally from the fact that contact with English was more recent, and less intensive, than contact with Sicilian and Italian.

Table 5: Composition of semantic fields by source language (Comrie & Spagnol 2016: 327)

Semantic field	Arabic	Romance	English	Misc.
Modern world	3.0%	65.3%	22.8%	9.0%
Animals	47.8%	29.1%	13.9%	9.1%
Clothing and grooming	38.7%	47.2%	10.4%	3.8%
Warfare and hunting	28.8%	65.0%	2.5%	3.8%
Law	36.0%	50.0%	0.0%	14.0%
Social and political relations	48.4%	48.4%	0.0%	3.2%

Unsurprisingly, English is best represented in the category of items relating to the modern world, but even here Romance dominates. Examples include English-derived *televixin* ‘television’ and Italian-derived *kafè* ‘coffee’.

The domain of animals divides rather neatly as follows. Common animals (especially land animals) of the Mediterranean area are largely Arabic-derived (e.g. *fenek* ‘rabbit’ < Maghrebi Arabic *fanak* ‘fennec fox’), while well-known non-indigenous animals are largely Romance-derived (e.g. *ljunfant* ‘elephant’ < Sic. *liufanti*, the additional /n/ perhaps the result of influence from *ljun* ‘lion’). More exotic animals, if there is a corresponding Maltese item at all, derive from English (e.g. *tapir* ‘tapir’). Clothing and grooming presents a similar picture, with Arabic-

derived *suf* ‘wool’, Sicilian-derived *ngwanta* ‘glove’, and English-derived *fer* ‘fur’, as does warfare and hunting, with Arabic-derived *sejf* ‘sword’, Sicilian-derived *xkubetta* ‘gun’, and English-derived *senter* ‘shotgun’ (< *centre-breech-loading shotgun*).

The total lack of English loans in the domains of law and social and political relations, at least in Comrie and Spagnol’s sample, is remarkable, given the extent to which the English language dominated public life in Malta in the twentieth century. A generalization that underlies this finding is that while English influence is strongest in the spheres of commerce, consumerism and, especially in the 21st century, popular culture (e.g. *vawċer* ‘voucher’, *ċċettja* ‘to chat’),¹⁸ at least as far as Maltese lexicon is concerned, it has not supplanted Italian in the domains of high culture and the affairs of state (e.g. *gvern* ‘government’ < Italian *governo*, *poeżija* ‘poem’ < Italian *poesia*).

3.4.2 Minor sources

Considering its location and the nature of population movements in the Mediterranean, it is hardly surprising that the Maltese lexicon also contains borrowings from languages other than Sicilian, Italian and English. The most obvious of these are borrowings from other Romance languages. First among them, as in other European languages, stands Latin, which provided a large chunk of Maltese scientific and technical vocabulary, whether as terminology (e.g. *ego*, *rektum* or *sukkursu* ‘underground water’), biological nomenclature (*fagu* ‘European beech, *Fagus sylvatica*’, *mirla* ‘brown wrasse, *Labrus merula*’) or set phrases and expressions (*ex cathedra*, *ibidem*). Curiously for a Catholic country, Latin is the source of very little religious vocabulary in Maltese; in this area, Maltese continues to rely almost exclusively on words of Arabic origin. Those Latin words related to religious matters employed in modern Maltese therefore typically refer to minutiae of Catholic Church rituals and procedures, such as *ekseat* ‘a bishop’s permission for a priest to leave the diocese’ (< *exeat*) or *indult* ‘a Pope’s authorization to perform an act otherwise not allowed by canon law’. Of the few Latin terms related to religion still in common use, *nobis* stands out as a rather curious lexical item: in Maltese, it is used as a (post-nominal) modifier indicating intensity or size, as

¹⁸Until at least 1991, when the Maltese government opened up television broadcasting rights to more than just the single state broadcaster TVM, Italian television stations, whose broadcasts from Sicily could be received in Malta, were very widely watched, and there was consequently considerable Italian influence on Maltese popular culture (Sammut 2007). This influence has waned considerably at the expense of English and American culture since the advent of broadcast pluralism in Malta, and especially with the rise of cable television and online video streaming.

Christopher Lucas & Slavomír Čéplö

in *tkaxkira nobis* ‘a sound thrashing’ or *tindifa nobis* ‘a thorough cleaning’.

Before the Order of Saint John gained control of Malta, the islands were for more than two centuries a part (whether officially or not) of the Crown of Aragon. As such, one would expect that speakers of Maltese during that era found themselves exposed to the languages of the Crown like Catalan, Spanish and Occitan, that and this was then reflected in the Maltese lexicon. In truth, however, there are only a few Maltese words that can clearly be traced to Ibero-Romance. [Biosca & Castellanos \(2017\)](#) identify a number of lexical items with Catalan and/or Occitan origins, but note that many of them can also be found in Sicilian, which in most cases can be clearly determined as the origin of the loan. On the other hand, there are Maltese words of obviously Romance origin whose current shape cannot be easily explained by any of the processes by which Sicilian or Italian words were made to conform to Maltese phonology, and where the Catalan or Occitan origin postulated by [Biosca & Castellanos \(2017\)](#) may offer a better explanation than that of “local formation” resorted to by previous works. These may include: *boxxla* ‘compass’ < Catalan *búixola* vs. Italian *bussola*; *frixa* ‘pancreas’ < Catalan *freixura* ‘entrails’ and even the very frequent *żgur* ‘certain’, which, due to its phonology, especially the /g/ (see §3.1.1.1), points to an origin in Catalan *segur* or Spanish *seguro*, rather than to its (Tuscan) Italian or Sicilian cognates, which both feature a /k/ in its place. These and other lexical items, onomastics (see [Biosca & Castellanos 2017](#): 46), and even usage (such as the ubiquitous Maltese swear word *l-ostja*, literally ‘the host, sacramental bread’, which is very atypical for Italian or Sicilian, but has a counterpart in the Spanish *la hostia*) suggest some influence of Ibero-Romance on Maltese which is yet to be thoroughly researched.

The much shorter French occupation of the Maltese islands left very little linguistic trace, and so it is internationalisms in the semantic field of culture (*bonton* ‘high society’, *etikett* ‘etiquette’), fashion (*manikin* ‘manequin’) and the culinary arts (*fundan* ‘fondant’, *ragu* ‘ragout’) where French borrowings in Maltese can be found. The few notable exceptions include *berġa* (< *auberge*), the term used for the residences of langues (chapters) of the Order of Saint John. The most prominent of these palaces, *Berġa ta’ Kastilja*, now houses the office of the Prime Minister of Malta, for which the term *Berġa* is often used metonymically. The other two Maltese words of French origin still in frequent daily use both happen to be connected to transportation: *xufier* (< *chauffeur*) ‘driver’ and *xarabank* (< *char à bancs*) ‘bus’. The latter is particularly interesting due to its pronunciation /ʃerɐˈbɛnk/, which indicates that it was borrowed directly from French and not from English (which would give /ʃerɐˈbɛnk/, as well as for its connection to the French-speaking Maghreb, where the same word was in use; this indicates the

possibility that it was brought from there by Maltese expatriates.

In addition to Romance languages, post-classical Greek, with its ubiquitous presence all across the Mediterranean (including the neighboring Sicily), could not help but leave a trace on Maltese vocabulary, small though it is. Aquilina (1976: 23) gives *Lapsi* ‘Feast of Ascension’ (< ἀνάληψη) as the solitary example of a Maltese religious term not inherited from Christian Arabic or borrowed from Romance languages. The other two examples of Greek loanwords involve a completely different sphere. The first is *ħamallu* ‘lewd, vulgar person’ < Greek χαμάλης (Dimitrakou 1958: 7781). This word may ultimately be traceable to Arabic (through Turkish), as is evident from its other meaning in Greek, namely ‘porter’ (< *ħammāl*). However, the meaning in which it appears in Maltese is unique to the Greek word, indicating that it was borrowed into Maltese from Greek. The other such term is *vroma* ‘complete failure, fiasco’ which is quite straightforwardly traceable to the Greek βρόμα/βρώμα ‘dirt, filth’ (Dimitrakou 1958: 1506, 1516).

With regard to the debates on the origin and history of Maltese, borrowings from other Afro-Asiatic languages have long been at the center of attention of Maltese etymological research. Berber is perhaps the most notorious example here, with a number of items cited as having Berber origins by Colin (1957) and Aquilina (1976: 25–39). Aquilina’s list is an expansion of Colin’s and thus both feature the same conspicuous items, which for the most part involve zoology, such as *fekruna* ‘tortoise’ (< *fekrun*, Naït-Zerrad 2002: 553) and *gendus* ‘bull’ (< *agenduz*, Naït-Zerrad 2002: 827). Additionally, Aquilina postulates a Berber origin for a number of lexical items where this seems questionable. In some cases the items in question are obviously Arabic loanwords in Berber (as with *bilħaq* ‘by the way’, quite transparently from Arabic *b-il-ħaq* ‘in truth’). In other cases subsequent research has argued against a Berber origin. For example, while Aquilina identifies *ženbil* ‘a large carrying basket’ as having a Berber origin, Borg (2004: 261) notes that it can also be found in the Arabic dialect of Aleppo and Arbil, and traces its ultimate origin to Akkadian through Aramaic. A large group of similarities between Maltese and Berber identified by Aquilina involve “Berber nursery language”, containing items like Berber *papa* ‘bread’ / Maltese *pappa*, Berber *ppspps* or *ppssi* ‘urine’ / Maltese *pixxa*, and Berber *kakka/qaqah* / Maltese *kakka* (both having to do with defecation). These forms are actually attested cross-linguistically (Ferguson 1964) at least as far north as Slovak (Ondráčková 2010) and cannot thus be considered loans from Berber. Nevertheless, the fact that there is a Berber lexical component in Maltese is well established, and Souag (2018) has shown that it may be larger than previously thought (e.g. his case for

the Berber etymology of the frequent adjective *ikejken* ‘small’).

Finally, in addition to Berber, Maltese also contains a small number of words that can be reasonably traced back to Aramaic. Along with obsolescent lexical items such as *ženbil* given above or *andar* ‘threshing floor’ (Behnstedt 2005: 116–117), this small list includes the frequent verb *xandara* ‘to broadcast, to spread (news)’, otherwise unattested in any other variety of Arabic (Borg 1996: 46). This verb is presumably derived from the common Aramaic root $\sqrt{\text{šdr}}$ ‘to dispatch, send’ with cognates in Mandaic (Drower & Macúch 1963: 450), Jewish Babylonian Aramaic (JBA; Sokoloff 2002: 1112–1113) and Christian Neo-Aramaic (Khan 2008: 1179). The insertion of [n] reflects the dissimilation of the geminated [dd] into [nd] (Lipiński 1997: 175–176); the same phenomenon involving the original geminated [bb] can also account for *ženbil* (cf. JBA *zabbīlā*; Sokoloff 2002: 397). These borrowings could on the one hand strengthen the case for a Levantine substrate in (if not origin of) Maltese, as Borg (1996) insists; on the other hand, some of them can also be found in other North African varieties (Behnstedt 2005).

4 Conclusion

This chapter has reviewed the extensive changes that have taken place in Maltese as a result of contact with Sicilian, Tuscan Italian, English, and other languages. The changes due to contact with Italo-Romance languages are so striking, especially but by no means only with respect to lexicon, that it is almost misleading to speak of these contacts having changed “Maltese”. Rather it might be argued that it was a Maghrebi Arabic dialect like any other that was subjected to these changes, and that Maltese, the distinct language that its speakers now feel it to be, was what emerged only once these changes were complete. The result is a language in which typically Semitic and typically Indo-European elements exist side-by-side at all linguistic levels.

The elements of contemporary standard Maltese that are the result of contact, summarized in this chapter, are now relatively well understood. But the language has naturally also evolved in numerous ways that owe little or nothing to the effects of contact with other languages. With a few notable exceptions (e.g. Borg 1978; Vanhove 1993), these changes have received far less attention. A desideratum for future historical linguistic work on Maltese is therefore to redress this imbalance.

Concerning contact-induced change specifically, future research could fruitfully include comparative work on the differential effects of contact on standard versus dialectal Maltese. And to the extent that it is possible, the field would

benefit greatly from a detailed history of the sociolinguistic effects of language contact in Malta in the early modern period.

Further reading

Krier (1976) is a short monograph on the influence of Italo-Romance on Maltese phonology, morphology, syntax, and lexicon.

Mifsud (1995) gives an in-depth description of Maltese loaned verbs.

Comrie & Spagnol (2016) examine lexical borrowing in Maltese in the context of loanword typology crosslinguistically.

Drewes (1994) and **Stolz (2003)** explore the question of whether Maltese is properly labeled a “mixed language”.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	L1, L2	1st, 2nd language
ADJ	adjective	M	masculine
BCE	before Common Era	MUDTv1	Maltese Universal Dependencies Treebank v1
BCv3	Bulbulistan corpus malti v3		
CE	Common Era	N	noun
COMP	complementizer	NEG	negative (particle)
DAT	dative	NP	noun phrase
DEF	definite article	OBJ	object
DEM	demonstrative	OBL	oblique
DEP	dependent form	PTCP	participle
DOM	differential object marking	PASS	passive
F	feminine	PL	plural
JBA	Jewish Babylonian Aramaic	PROG	progressive
GEN	genitive	PRF	perfect (suffix conjugation)
IMPF	imperfect (prefix conjugation)	RL	recipient language
INF	infinitive	Sic.	sicilian
It.	Italian	SG	singular
PRG	pragmatic marker	SING	singulative
		SL	source language

Christopher Lucas & Slavomír Čéplö

Acknowledgments

The research presented in this chapter was partly funded by a Leadership Fellows grant from the UK Arts and Humanities Research Council, grant number APVV-15-0030 from The Slovak Research and Development Agency (APVV), and the ERC Starting Grant number 679083 from the European Research Council, whose support is hereby gratefully acknowledged.

Primary sources

Maltese examples above are primarily cited from the general corpus of Maltese *bulbulistan corpus mali v3* (accessible at www.bulbul.sk/bonito2, login: guest, password: Ghilm3), as well as from the *Maltese Universal Dependencies Treebank v1* (accessible at www.bulbul.sk/annis-gui-3.4.4/), both described as to their composition and annotation in Čéplö (2018). Each citation is accompanied by an abbreviation identifying the source (BCv3 and MUDTv1, respectively), as well as the specific document where it can be found.

References

- A descriptive and comparative grammar of Andalusí Arabic*. 2013. Leiden: Brill.
- Aquilina, Joseph. 1976. *Maltese linguistic surveys*. Malta: The University of Malta.
- Avram, Andrei A. 2012. Some phonological changes in maltese reflected in onomastics. *Bucharest Working Papers in Linguistics* (14). 99–119.
- Avram, Andrei A. 2014. The fate of the interdental fricatives in maltese. *Romano-Arabica* (14). 19–32.
- Avram, Andrei A. 2017. Word-final obstruent devoicing in maltese: inherited, internal development or contact-induced? *Academic Journal of Modern Philology* 6. 23–37.
- Ballou, Maturin M. 1893. *The story of malta*. Cambridge, MA: The Riverside Press.
- Behnstedt, Peter. 2005. Voces de origen sirio y yemení en el árabe magrebí / andalusí. In Jordi Aguadé, Ángeles Vicente & Leila Abu Shams (eds.), *Sacrum Arabo-Semiticum*, 115–122. Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo.
- Biosca, Carles & Carles Castellanos. 2017. Aspects of the comparison between Maltese, Mediterranean Lingua Franca and the Occitan–Catalan linguistic group. In Benjamin Saade & Mauro Tosco (eds.), *Advances in Maltese linguistics*, 39–64. Berlin: De Gruyter.

- Blau, Joshua. 1967. *A grammar of Christian Arabic based mainly on South-Palestinian texts from the first millennium, fasc. ii: §§170–368*. Louvain: Secrétariat du CorpusSCO.
- Blouet, Brian. 1967. *The story of malta*. London: Faber & Faber.
- Borg, Albert. 1987–88. To be or not to be: a copula in Maltese. *Journal of Maltese Studies* 17–18. 54–71.
- Borg, Albert & Marie Azzopardi-Alexander. 1997. *Maltese*. London: Routledge.
- Borg, Albert & Michael Spagnol. 2015. Nominal sentences in Maltese: a corpus-based study. Paper presented at the 5th International Conference on Maltese Linguistics, L'Università di Torino, June 24–26.
- Borg, Alexander. 1978. *A historical and comparative phonology and morphology of maltese*. Jerusalem: The Hebrew University of Jerusalem doctoral dissertation.
- Borg, Alexander. 1985. *Cypriot Arabic: a historical and comparative investigation into the phonology and morphology of the arabic vernacular spoken by the Maronites of Kormakiti village in the Kyrenia district of north-west Cyprus*. Stuttgart: Deutsche Morgenländische Gesellschaft.
- Borg, Alexander. 1994. Some evolutionary parallels and divergences in cypriot arabic and maltese. *Mediterranean Language Review* 8. 41–67.
- Borg, Alexander. 1996. On some levantine linguistic traits in maltese. *Israel Oriental Studies* 16. 133–152.
- Borg, Alexander. 1997. Maltese phonology. In Alan S. Kaye (ed.), *Phonologies of asia and africa*, vol. 1, 245–285. Winona Lake, IN: Eisenbrauns.
- Borg, Alexander. 2004. *A comparative glossary of Cypriot Maronite Arabic (Arabic–English): With an introductory essay*. Leiden: Brill.
- Bovingdon, R. & A. Dalli. 2006. Statistical analysis of the source origin of maltese. In A. Wilson, D. Archer & P. Rayson (eds.), *Corpus linguistics around the world*, 63–76. Amsterdam: Rodopi.
- Brincat, Joseph M. 1995. *Malta 870–1054. al-Ḥimyarī's account and its linguistic implications*. Valetta: Said International.
- Brincat, Joseph M. 1996. Maltese words. An etymological analysis of the Maltese lexicon. In Jens Lüdtke (ed.), *Romania Arabica. Festschrift für Reinhold Kontzi zum 70. Geburtstag*, 111–115. Tübingen: Gunter Narr.
- Brincat, Joseph M. 2011. *Maltese and other languages. a linguistic history of Malta*. Malta: Midsea Books.
- Brincat, Joseph M. & Manwel Mifsud. 2015. Maltese. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: an international handbook of the languages of europe*, 3349–3366. Berlin: De Gruyter Mouton.

Christopher Lucas & Slavomír Čéplö

- Camilleri, Maris. 2011. On pronominal verbal enclitics in Maltese. In Sandro Caruana, Ray Fabri & Thomas Stolz (eds.), *Variation and change: the dynamics of Maltese in space, time and society*, 131–156. Berlin: Akademie Verlag.
- Camilleri, Maris & Louisa Sadler. 2018. The grammaticalisation of a stage-level copula in vernacular Arabic. Paper presented at the 32nd Annual Symposium on Arabic Linguistics, Arizona State University, February 23–25.
- Caubet, Dominique. 2011. Moroccan Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Census of Population and Housing 2011: Final Report*. 2014. Tech. rep. Valletta: National Statistics Office.
- Čéplö, Slavomír. 2014. An overview of object reduplication in Maltese. In Albert Borg, Sandro Caruana & Alexandra Vella (eds.), *Perspectives on Maltese linguistics*, 201–222. Berlin: De Gruyter.
- Čéplö, Slavomír. 2018. *Constituent order in Maltese: a quantitative analysis*. Prague: Charles University doctoral dissertation.
- Coghill, Eleanor. 2014. Differential object marking in Neo-Aramaic. *Linguistics* 52(2). 335–364.
- Cohen, David. 1966. Le système phonologique du maltais: aspects synchroniques et diachroniques. *Journal of Maltese Studies* 3. 1–26.
- Cohen, David. 1970. Le système phonologique du maltais: aspects synchroniques et diachroniques. In David Cohen (ed.), *Etudes de linguistique arabe et sémitique*, 126–149. The Hague: Mouton.
- Colin, Georges Séraphin. 1957. Mots «berbères» dans le dialecte arabe de Malte. In *Mémorial André Basset*, 7–16. Paris: Adrien Maisonneuve.
- Comrie, Bernard & Michael Spagnol. 2016. Maltese loanword typology. In Gilbert Puech & Benjamin Saade (eds.), *Shifts and patterns in Maltese*, 315–329. Berlin: De Gruyter Mouton.
- de Cat, Cécile. 2010. *French dislocation. Interpretation, syntax, acquisition*. Oxford: Oxford University Press.
- de Sacy, Silvestre. 1829. [no title]. *Journal des Savants* 4. 195–204.
- de Soldanis, Giovanni Pietro Francesco Agius. 1750. *Della lingua punica presentemente usata da maltesi*. Rome: Generoso Salomoni.
- Dimitrakou, Dimitriou. 1958. *Méga leksikón tés Ellénikés Glósses*. Athens: Ekdoseis Domi.
- Döhla, Hans-Jörg. 2016. The origin of differential object marking in Maltese. In Gilbert Puech & Benjamin Saade (eds.), *Shifts and patterns in Maltese*, 149–174. Berlin: De Gruyter Mouton.

- Drewes, A. J. 1994. Borrowing in maltese. In Peter Bakker & Maarten Mous (eds.), *Mixed languages. 15 case studies in language intertwining*. Amsterdam: Institute for Functional Research into Language & Language Use.
- Drower, E.S. & R. Macúch. 1963. *A Mandaic dictionary*. Clarendon Press.
- Fabri, Ray. 1993. *Kongruenz und die Grammatik des Maltesischen*. Tübingen: Max Niemeyer.
- Fabri, Ray & Albert Borg. 2002. Topic, focus and word order in Maltese. In Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi & Zakia Iraqui-Sinaceur (eds.), *Aspects of the dialects of arabic today*, 354–363. Rabat: AMAPATRIL.
- Fenech, Edward. 1978. *Contemporary journalistic Maltese*. Leiden: E. J. Brill.
- Ferguson, Charles A. 1964. Baby talk in six languages. *American Anthropologist* 66(6). 103–114.
- Fiorini, Stanley. 1986. The resettlement of gozo after 1551. *Melita Historica* (9). 203–244.
- Flege, James E., Murray J. Munro & I.R.A. MacKay. 1995. Effects of age of second-language learning on the production of english consonants. *Speech Communication* 16. 1–26.
- Friedman, Victor A. 2008. Balkan object reduplication in areal and dialectological perspective. In *Clitic doubling in the Balkan languages*, 35–63. Amsterdam: John Benjamins.
- Gabriel, Christoph & Elena Kireva. 2014. Prosodic transfer in learner and contact varieties: speech rhythm and intonation of buenos aires spanish and l2 castilian spanish produced by italian native speakers. *Studies in Second Language Acquisition* 36. 257–281.
- Gardani, Francesco. 2012. Plural across inflection and derivation, fusion and agglutination. In Lars Johanson & Martine Robbeets (eds.), *Copies versus cognates in bound morphology*, 71–97. Leiden: Brill.
- Gatt, Albert. 2018. Definiteness agreement and the pragmatics of reference in the Maltese np. *STUF – Language Typology and Universals* 71(2). 175–198.
- Gatt, Albert & Ray Fabri. 2018. Borrowed affixes and morphological productivity: a case study of two maltese nominalisations. In Patrizia Paggio & Albert Gatt (eds.), *The languages of malta*, 143–169. Berlin: Language Science Press.
- Geary, Jonathan. 2017. The historical development of the maltese plural suffixes *-iet* and *(i)jiet*. Paper presented at the 23rd International Conference on Historical Linguistics, San Antonio, TX, 31 July – 4 August.
- Gesenius, Wilhelm. 1810. *Versuch über die maltesische Sprache zur Beurtheilung der neulich wiederhohltten Behauptung, dass sie Ueberrest der altpunischen sey, und als Beytrag zur arabischen Dialektologie*. Leipzig: Fr. Chr. Wilh. Vogel.

Christopher Lucas & Slavomír Čéplö

- Gibson, Maik. 2011. Tunis Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Goodwin, Stefan. 2002. *Malta: mediterranean bridge*. Westport, CT: Bergin & Garvey.
- Grech, Paul. 1961. Are there any traces of Punic in Maltese? *Journal of Maltese Studies* 1. 130–138.
- Grice, Martine, Alexandra Vella & Anna Bruggeman. 2019. Stress, pitch accent, and beyond: intonation in maltese questions. *Journal of Phonetics* 76. 100913.
- Harrell, Richard S. 2004. *A short reference grammar of Moroccan Arabic*. Washington, DC: Georgetown University Press.
- Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the world's languages. a comparative handbook*. Berlin: De Gruyter Mouton.
- Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London: Routledge.
- Herin, Bruno & Martin R. Zammit. 2017. Three for the price of one: the dialects of kerkennah (tunisia). In Veronika Ritt-Benmimoun (ed.), *Tunisian and libyan arabic dialects: common trends – recent developments – diachronic aspects*, 135–146. Zaragoza: Prensas de la Universidad de Zaragoza.
- Iemmolo, Giorgio. 2010. Topicality and differential object marking. Evidence from Romance and beyond. *Studies in Language* 34(2). 239–272.
- Khan, Geoffrey. 1984. Object markers and agreement pronouns in Semitic languages. *Bulletin of the School of Oriental and African Studies* 43(3). 468–500.
- Khan, Geoffrey. 2008. *The Neo-Aramaic dialect of Barwar*. Leiden: Brill.
- Klimiuk, Maciej. 2018. Preliminary remarks on the gozitan dialect of gharb, malta. Paper presented at the 12th conference of the Association Internationale de Dialectologie Arabe, Marseille, 30 May – 2 June.
- Krapova, Iliyana & Guglielmo Cinque. 2008. Clitic reduplication constructions in Bulgarian. In Dalina Kallulli & Liliane Tasmowski (eds.), *Clitic doubling in the Balkan languages*, 257–287. Amsterdam: John Benjamins.
- Krier, Fernande. 1976. *Le maltais au contact de l'italien. etude phonologique, grammaticale et sémantique*. Hamburg: Helmut Buske.
- Lahdo, Ablahad. 2009. *The Arabic dialect of Tillo in the region of Siirt (south-eastern Turkey)*. *Acta Universitatis Upsaliensis*. Uppsala: Uppsala University.
- Lipiński, Edward. 1997. *Semitic languages. Outline of a comparative grammar*. Leuven: Peeters.
- Loporcaro, Michele. 2011. Phonological processes. In Martin Maiden, John Charles Smith & Adam Ledgeway (eds.), *The cambridge history of the romance languages*, vol. 1: *Structures*, 109–154. Cambridge: Cambridge University Press.

- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Maiden, Martin & Cecilia Robustelli. 2007. *A reference grammar of modern italian*. London: Routledge.
- Mifsud, Manwel. 1995. *Loan verbs in Maltese: a descriptive and comparative study*. Leiden: Brill.
- Mifsud, Manwel. 2011. Maltese. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Mifsud, Manwel & Albert Borg. 1997. *Fuq l-Għatba tal-Malti: deskrizzjoni komunikattiva ta' l-ilsien Malti fuq il-mudell tat-Threshold Level 1990*. Council of Europe.
- Naït-Zerrad, Kamal. 2002. *Dictionnaire des racines berbères (formes attestées), III. D–GEY*. Leuven: Peeters.
- O'Rourke, Erin. 2005. *Intonation and language contact: a case study of two varieties of peruvian spanish*. Urbana-Champaign: University of Illinois doctoral dissertation.
- Ondráčková, Zuzana. 2010. *Komparatívny výskum detskej lexiky*. Prešov: Prešovská univerzita.
- Piccitto, Giorgio (ed.). 1977. *Vocabolario siciliano*. Vol. 1: (A-E). Catania: Centro di Studi Filologici e Linguistici Siciliani.
- Ritt-Benmimoun, Veronika. 2014. *Grammatik des arabischen Beduinendialekts der Region Douz (Südtunesien)*. Wiesbaden: Harrassowitz.
- Russi, Cinzia. 2008. *Italian clitics: An empirical study*. Berlin: Mouton de Gruyter.
- Saade, Benjamin. 2019. Assessing productivity in contact: italian derivation in maltese. *Linguistics* 57(1).
- Sammut, Carmen. 2007. *Media and maltese society*. Plymouth: Lexington Books.
- Sciriha, Lydia & Mario Vassallo. 2006. *Living languages in malta*. Malta: Print IT Printing Services.
- Seifart, Frank. 2015. Direct and indirect affix borrowing. *Language* 91. 511–532.
- Seifart, Frank. 2017. Patterns of affix borrowing in a sample of 100 languages. *Journal of Historical Linguistics* 7. 389–431.
- Sokoloff, Michael. 2002. *A dictionary of Jewish Babylonian Aramaic of the Talmudic and Geonic periods*. Ramat Gan: Bar Ilan University Press.
- Souag, Lameen. 2013. *Berber and arabic in siwa (egypt): a study in linguistic contact*. Cologne: Rüdiger Köppe.
- Souag, Lameen. 2017. Clitic doubling and language contact in Arabic. *Zeitschrift für Arabische Linguistik* 66. 45–70.

Christopher Lucas & Slavomír Čéplö

- Souag, Lameen. 2018. Berber etymologies in Maltese. *International Journal of Arabic Linguistics* 4(1). 190–223.
- Spagnol, Michael. 2011. *A tale of two morphologies. Verb structure and argument alternations in Maltese*. Konstanz: Universität Konstanz doctoral dissertation.
- Stolz, Thomas. 2003. Not quite the right mixture: Chamorro and Malti as candidates for the status of mixed language. In Yaron Matras & Peter Bakker (eds.), *The mixed language debate: theoretical and empirical advances*, 271–315. Berlin: De Gruyter Mouton.
- Sutcliffe, Edmund. 1936. *A grammar of the Maltese language. With chrestomathy and vocabulary*. Oxford: Oxford University Press.
- Taine-Cheikh, Catherine. 2011. Numerals. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Ullmann, Manfred. 1989. *Adiminiculum zur Grammatik des klassischen Arabisch*. Wiesbaden: Harrassowitz.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Vanhove, Martine. 1993. *La langue maltaise: etudes syntaxiques d'un dialecte arabe «périphérique»*. Wiesbaden: Harrassowitz.
- Vanhove, Martine. 1998. De quelques traits pré-hilaliens en maltais. In Jordi Aguadé, Patrice Cressier & Ángeles Vicente (eds.), *Peuplement et arabisation au Maghreb occidental: dialectologie et histoire*, 97–108. Zaragoza: Casa de Velásquez.
- Vassalli, Michelantonio. 1791. *Mylsen phoenico-punicum sive grammatica melitensis*. Rome: Antonio Fulgoni.
- Vella, Alexandra. 1994. *Prosodic structure and intonation in maltese and its influence on maltese english*. Edinburgh: University of Edinburgh doctoral dissertation.
- Vella, Alexandra. 2003. Language contact and maltese intonation: some parallels with other language varieties. In Kurt Braunmüller & Gisella Ferraresi (eds.), *Aspects of multilingualism in european language history*, 261–283. Amsterdam: John Benjamins.
- Vella, Alexandra. 2009. On maltese prosody. In Bernard Comrie, Ray Fabri, Elizabeth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing maltese linguistics*, 47–68. Amsterdam: John Benjamins.
- Versteegh, Kees. 2014. Pidgin verbs: infinitives or imperatives. In Isabelle Buchstaller, Anders Holmberg & Mohammad Almoaily (eds.), *Pidgins and creoles beyond africa-europe encounters*. Amsterdam: John Benjamins.

- Woidich, Manfred. 2006. *Das kairenisch-arabische: eine grammatik*. Wiesbaden: Harrassowitz.
- Zagona, Karen. 2002. *The syntax of Spanish*. Cambridge: Cambridge University Press.
- Zammit, Martin R. 2014. The Sfaxi (tunisian) element in maltese. In Albert Borg, Sandro Caruana & Alexandra Vella (eds.), *Perspectives on maltese linguistics*, 23–44. Berlin: Akademie Verlag.

Chapter 7

Nigerian Arabic

Jonathan Owens

University of Bayreuth

Nigerian Arabic displays an interesting interplay of maintenance of inherited structures along with striking contact-induced innovations in a number of domains. This chapter summarizes the various domains where contact-based change has occurred, concentrating on those less studied not only in Arabic linguistics, but in linguistics in general, namely idiomatic structure and an expanded functionalization of demonstratives. Methodologically, comparative corpora are employed to demonstrate the degree of contact-based influence.

1 Historical and linguistic background

Nigerian Arabic (NA) is spoken by perhaps – there are no reliable demographic figures from the last 50 years – 500,000 speakers. These are found mainly in northeast Nigeria in the state of Borno where their homeland is concentrated along the Cameroon–Chad border as far south as Banki, spreading westwards towards Gubio, and south of Maiduguri towards Damboa. Mirroring a larger trend in Nigerian demographics, the past 40 years have seen a considerable degree of rural–urban migration. This has seen, above all, the development of large Arab communities in cities in Borno – the capital Maiduguri has at least 50,000 alone¹ – though they are now found throughout cities in Nigeria.

Arabs in Nigeria are traditionally cattle nomads, part of what the anthropologist Ulrich Braukämper (1994) has called the “Baggara belt”, named after the

¹A report in the 1970s by an urban planning company, the Max Lock Group (1976) estimated that 10% of the then estimated population of 200,000 Maidugurians were Arabs. Today the population of Maiduguri is not less than one million and may be considerably larger, which proportionally would estimate an Arab population in Maiduguri alone of at least 100,000. Of course, if one included the refugee camps today, the number would be much higher.



Jonathan Owens

Arab tribe in the western Sudan (Kordofan, Darfur; see [Manfredi 2010](#)) whose culture and dialect are very similar to those of the Nigerian Arabs. Until the very recent Bokko Haram tragedy, besides nomadism, Arabs practiced subsistence farming. As of the writing of this chapter, nearly all rural Nigerian Arabs have been forced to flee their home villages and cattle camps and are living mainly in refugee camps in northeast Nigeria and neighboring countries.

Arabs first came to the Lake Chad area – whether territorial Nigeria is at this point undetermined – in the late fourteenth century. They were part of what initially was a slow migration out of Upper Egypt towards the northern Sudan beginning in the early thirteenth century, which gained momentum after the fall of the northern Nubian kingdom of Nobadia (or Maris) in the fourteenth century. All in all, NA exhibits a series of significant isoglosses which link it to Upper Egypt, via Sudanese Arabic, even if it displays interesting “archaisms” linking it to regions far removed from Africa ([Owens 2013](#)). Its immediate congeners are found in what I have termed Western Sudanic Arabic (WSA; [Owens 1994a,b](#)), stretching between northeast Nigeria in the west and Kordofan in the east ([Manfredi 2010](#)). When properties of NA are contrasted with other varieties of Arabic, it is implicitly understood that these do not necessarily include other WSA varieties. Much more empirical work is necessary in this regard, but, to give one example, many of the extended functions of the NA demonstrative described in §3.3.2 below are also found in Kordofanian Arabic ([Manfredi 2014](#)). Moreover, where throughout the Sudanic region as a whole any given isogloss lies is also an open question, as is the issue of the degree to which the contact-induced changes suggested here represent broad areal phenomena. As my own in many cases detailed data derives from NA, I limit most observations to this area. NA itself divides into two dialect areas, a western and an eastern one that I have also termed Bagirmi Arabic, since it is spoken by Arabs in the Bagirmi-speaking region.

In Borno, Arabs are probably the largest minority ethnic group, though still a minority. The entire area bordering Lake Chad, both to the east and to the west, is dominated by Kanuri-speaking peoples (Kanembu in Chad). This was a domination which the Arabs already met in their first migrations into the region, both a political and a linguistic domination. As will be seen, this has left dramatic influences in some domains of NA, while leaving others untouched.

While until about 1970 Kanuri was the dominant co-territorial language, Arabs in the Lake Chad area have been in close contact with other languages/ethnic groups as well, for instance Fulfulde, Kotoko (just south of Lake Chad) and Bagirmi (south of Ndjamena in Chad). Furthermore, Kanuri established itself in Borno

in an area already populated by speakers of Chadic languages, so it as well was probably influenced by some of the co-territorial languages Arabs met. Since 1970, Hausa has become the dominant lingua franca in all urban areas in north-eastern Nigeria (indeed throughout the north of the country). In a sample of 58 Maiduguri speakers for instance (Owens 1998; Owens 2000: 324), 50 professed knowing Hausa, and 46 Kanuri. In the only study of its type, Broß (2007) shows that urban Maiduguri Nigerian Arabs have a high degree of accuracy for a number of complex variables in Hausa, while, using a similar sample, in one of the few interactional studies available, Owens (2002) also documents a high multilingual proficiency between Arabic and Hausa, and for some speakers, English. How such micro-studies can be interpreted against the over 400 years of NA contact with area languages remains a question for the future. Rural areas have not yet experienced such a high penetration of Hausa. In a second, rural sample consisting of 48 individuals, only sixteen self-reported knowing Hausa versus forty Kanuri. Note that as of the 1990's, there were still a considerable number of monolingual Arabic speakers, particularly in the area along the Cameroonian border which among Nigerian Arabs is known as the Kala-Balge region.

While Standard Arabic (Classical Arabic) has always been a variety known among a small educated elite in Borno (of all ethnic backgrounds), along with Hausa it has gained considerable momentum in recent years. Whereas traditionally Classical Arabic, as a part of Koranic memorization, has always been a part of Arabs' linguistic repertoire, it is only since about 1990 that the teaching of Standard Arabic as a school subject has spread oral fluency in this variety.

To this point, conditions have been described which, on paper at least, would favor influence via borrowing under RL-agentivity (in the terminology of Van Coetsem 1988; 2000). Nigerian Arabs as a linguistic minority tend to be bilingual, and, it may be assumed, have had a history of bilingualism in Kanuri and locally other languages going back to their first migrations into the region. Equally, however, Nigerian Arabic society has itself integrated other ethnic groups creating conditions of shift to Arabic. According to Braukämper's (1994) thesis, the very basis of Nigerian Arab nomadism is cattle nomadism based on a Fulani model. This is said to have arisen around the mid-seventeenth century as Arabs coming from the east met Fulani moving west. Today there is very little Fulfulde spoken in Borno or Chad, so it may be surmised that the result of the Fulani-Arab contact was language shift in favor of Arabic. Furthermore, slavery was a well-established institution which incorporated speakers from other ethnic groups (see recording TV57b-Mule-Hawa in Owens & Hassan 2011–17, as an instance of a slave descendant). Intermarriage is another mechanism by which L1 speakers

Jonathan Owens

would switch to Arabic. In contemporary Nigeria, intermarriage in fact tends to favor Arab women marrying outside their group, rather than marriage into Arab society, though there is no cultural proscription of the practice, and such practices tend, *inter alia*, to be influenced by the relative prestige and power of the groups involved. Today Arabs are dominated politically by the Kanuri, though there are eras, for example the period of Kanemi in the mid-nineteenth century, or the rule of Rabeh at the beginning of the twentieth, when Arabs were more dominant and perhaps had greater access to marriage from outside groups. I will return to these summaries in §4.

The data for this chapter comes from long years of working on Arabic in the Lake Chad region. More concretely, a large oral corpus of about 400,000 words (Owens & Hassan 2011–17) forms the basis of much of the research, and this corpus will be referred to in a number of places in the chapter. When a form is said to be rare, frequent, etc., these evaluations are made relative to what can be found in the corpus. All examples come from this corpus. The source of the recording in the data bank is indicated by the number in brackets at the end of the example.

2 Contact and historical linguistics

Language contact is an integral part of historical linguistics. In the case of Arabic, the history of Arabic has different interpretations, so it is relevant here to very briefly reiterate my own views (Owens 2006). All varieties of contemporary Arabic derive from a reconstructed ancestor or ancestors. Whether singular or plural is a crucial matter, but one answered legitimately only within historical linguistic methodology (see e.g. Retsö 2013, who appears to favour the plural). As is usually accepted (perhaps not by some working within grammaticalization theory, e.g. Heine & Kuteva 2011), historical linguistics operates at the juncture of inheritance and contact, and examines change due to internal developments and change due to contact. In the case of Arabic, contact extends well into the pre-Islamic era (Owens 2013; 2016a; forthcoming).

Furthermore, it operates at the level of the speech community, and Arabic has and had many speech communities, each with its own linguistic history. The history of speech communities is not co-terminous with political history, usually not with the history of individual countries, or even with cultural entities such as a nomadic life style. It follows that Arabic linguistic history is quite complicated, its large population being the product of and reflecting many individual social entities.

Any individual contemporary Arabic speech community therefore lies at the end of many influences. Interpreting whether and when a particular change occurred due to contact is anything but straightforward, as I will discuss very briefly in the following phonological issue.

Ostensibly NA shows the loss of *θ:

- (1) *θ > t, *θawr > *tōr* ‘bull’

or in the eastern area:

- (2) *θ > s, *θawr > *sōr* ‘bull’

There is no space to go into the detailed historical linguistic arguments here, but it would be incorrect to assert that these changes, quite plausibly originally due to contact, took place in the territorial NA or WSA region. This can be seen *inter alia* in the fact that all of Egyptian Arabic (EA) and all of the Sudanic region including the WSA area has (1). Whenever the shift occurred, it was well before Arabs came to the Sudanic region, let alone Nigeria. The changes in (1) and, I would argue, (2) as well, are part of the historical linguistics of ancestral Sudanic Arabic, but the changes themselves are antecedent to Arabic in the Sudanic region and therefore are not treated here.

3 Contact-induced changes

3.1 Phonology

Excluding cases like (1)–(2) on methodological grounds, other than marginal effects due to borrowing, discussed briefly in §3.2, there are no significant instances of contact-induced phonological change limited only to NA. Two changes confined to all or part of the WSA region can be suspected, however.

Throughout Nigeria, Cameroon and most of Chadian Arabic *ḥ/ʕ have depharyngealized.

- (3) *ḥ/ʕ > h/?
ḥilim ‘dream’ > *hilim*
gaʕad ‘stay, sit’ > *gaʔad*

As a set, the change is attested only in this region. Moreover, the area it is attested in begins by and large in the region where Arabic fades into minority status.

Jonathan Owens

A second candidate for a local WSA innovation is the reflex of *t̥, which is a voiced, emphatic implosive /d/. The implosive /d/ is also found in Fulfulde, as well as other possible contact languages such as Bagirmi, which, as noted above, are one source of shifters to Arabic. Manfredi (2010: 44; and personal communication) notes that /d/ is an allophonic variant in Kordofanian Baggaara Arabic.

The status of one phoneme, /č/, is still open. It is fairly frequent (about 100 entries out of about 8,500 (excluding proper names) in a dictionary currently in preparation begin with /č/). In a minority of cases an Arabic origin is certain or likely, e.g. *čāl* ‘come’ (eastern variant) < *tāl and perhaps *čatt* ‘all’, < *šattā ‘various’, with [š + t] > /č/ recalling some Gulf dialects *ičūf* ‘you see’. /č/ is never a reflex of *k. However, most instances of /č/ are still unaccounted for (e.g. *ču* ‘very red’, *čāqab* ‘wade through’).

All in all then, there has not been a great deal of fundamental phonological change due to contact. Note that NA maintains all inherited emphatics, and probably inherited its phonemically contrastive emphatic /ṁ/, /ṛ/ and perhaps its /l̥/ as well.

3.2 Loanwords

Despite its long period as a minority language in the Lake Chad region, NA has only a modest number of loanwords (see Owens 2000 for a much more detailed treatment of all aspects of loanwords in the classical sense). In a token count based on about 500,000 words, only about 3% of all words were loans. On a type basis the percentage rises considerably, though still is far from overwhelming. Table 1 presents loanword provenance data from the dictionary currently in progress.

Table 1: Loanwords in NA, types, N = 1263

Language	Types
English	509
Hausa	255
Kanuri	252
Standard Arabic	212
French	21
Fulfulde	12
Kotoko	2

The figures in Table 1 are probably a slight underestimation, as there are about sixty words, like *bazingir* ‘soldier of Rabeh’ which clearly are not of Arabic origin but whose precise origin has not been found.

There are many interesting issues in understanding the loanwords, a few of which I mention very cursorily here. The semantic domains differ from source to source. Standard Arabic, for instance, has mainly learned words. Kanuri covers a fairly wide spectrum, and strikingly includes a large number of discourse markers and conjunctions, on a token basis. *dugó* ‘then, so’ (< *dugó*) for instance has something in the range of 630 occurrences and *yó, yō, iyō* ‘so, okay, aha’ has 938. In Owens (2000: 303), discourse particles and conjunctions are shown to make up no less than 23.3% of all loanword tokens in the sample. It is noticeable that although a few Hausa discourse marker tokens (*to* ‘right, okay, so’) do occur, there are hardly twenty in all, this being indicative of the much shorter time span Hausa has been in large-scale contact with NA as compared to Kanuri.

The question of origin has two aspects, one the ultimate origin, the other how it got into NA. *bel* ‘belt’ is ultimately of English origin, but the same word is also found in Hausa (*bel*) and in Kanuri (*bêl*). Given that both of these languages are dominant ones, it is likely that *bel* entered NA from one of these, not directly from English. The statistics above are the ultimate origin. The medial origin (travel words) is much harder to trace. Using the corpus, it is possible to discern likely paths. For instance, NA *sanāʔa* ~ *sapa* ‘trade, occupation, profession’ is cognate with both Standard Arabic *šināʔa* ‘art, occupation, craft’ and Hausa *sanāʔā* ‘trade, craft, profession’. Considering the distribution of *sapa* among speakers who have no knowledge of Standard Arabic, it is likely that the word reached NA via Hausa.

Non-Arabic phonology will often be maintained in the loanword. However, as can be discerned from loanwords of higher frequency, usually there is variation between retention of the source phoneme and adaptation. For instance ‘police’ comes in two forms, *polīs* and *folīs* (Owens 2000: 278). The [p] variant occurs in 19 tokens distributed among eight speakers, the [f] in 18 tokens among six speakers. Inspection of the statistics shows only a tendential bias towards [f] among women and villagers. Both variants appear therefore to be widespread. Note in this case that variation between [p] and [f] is also endemic to Kanuri, so it is likely here that the variation itself was borrowed.

3.3 Syntax

There are three strong candidates for contact-induced change in the syntactic domain: word order, ideophones and an expansion and realignment of demonstratives.

Jonathan Owens

3.3.1 Word order and ideophones

NA has only two pre-noun modifiers, *gōlit* ‘each’, *kunni* ‘each’.

- (4) *gōlit* ʔid *nulummu*
 each holiday gather.IMPF.1PL
 ‘We would gather at each festival.’

Otherwise NA is head-N-initial, which means that ‘all’ *čatt* and ‘how many’ *kam* are post-N, while demonstratives only have a post-N position (as in EA).

- (5) *numšu* *be* ʔaxuwāt-na *čatt*-ina
 go.IMPF.1PL with sisters-1PL all-1PL
 ‘We go with our sisters, all of us.’
- (6) *tağīb* *ḍahaḥ kam*
 bring.IMPF.2SG.M gold how.much
 ‘How much gold do you bring?’

The post-nominal-only demonstrative would have been inherited from EA. *čatt* ‘all’ mirrors the post-nominal alternative for *kull*, both taking a pronoun cross-referencing the head noun. Therefore, strictly speaking, the only innovation is the post-nominal position of *kam* ‘how many’, and an argument could be made that internal analogies lead NA towards a more consistent head-first noun-phrase order. By the same token, Kanuri is also consistently head-first order in the NP, so it could be that contact with Kanuri accelerated an inherited trend.

The numeral phrase has undergone considerable re-structuring. From ‘twenty’ upwards, the order is decade–ones.

- (7) *talātīn haw wāhid*
 thirty and one
 ‘thirty one’

Though inherited teens do occur, the usual structure is ten–ones.

- (8) ʔasara haw wāhid
 ten and one
 ‘eleven’

This order mirrors that of Kanuri (Hutchison 1981: 203), and indeed that of most languages in the immediate Lake Chad area. Uzbekistan Arabic has the

same numeral order and structure as NA, and in these cases independent contact events are likely the reason for the shift from an inherited structure.

A new syntactic category (for Arabic), that of ideophones, is described in detail in Owens & Hassan (2004) (see *tul* in (11b) below). To date in the dictionary of NA in progress there are 342 ideophones, about 4% of the lemma total.

3.3.2 Demonstratives

Formally, NA demonstratives reproduce their inherited forms, and therefore are virtually identical to paradigms found in various Egyptian dialects, except that, in consonance with NA morphology, feminine plural has a distinct form, which most Egyptian dialects have neutralized (see Table 2).

Table 2: NA demonstratives

	Near		Far	
	Singular	Plural	Singular	Plural
Masculine	da	dōl	ḏāk	dōlak
Feminine	di	dēl	ḏīk	dēlak

As with all Arabic demonstratives, NA demonstratives are used both as modifiers and pronominally. The traditional, inherited functions are entity referential (*al-bēt da* ‘this house’), and propositional anaphoric (*ʔašān da* ‘because of this’, where ‘this’ references an introduced proposition).

Additionally, however, the demonstratives occur in a number of contexts which either are not attested at all, or are attested only on an extremely infrequent basis in other Arabic dialects. I summarize these here.

- 1. Marking the end of dependent clauses, whether relative, conditional or adverbial.

Usually *da* is the default form in this function, though in the case of relative clauses the demonstratives often agrees with the head noun.

- (9) Conditional clause
[kan gul balkallam kalām-hum da] ma
[if say.PRF.1SG speak.IMP.1SG language-3PL.M DEM.SG.M] NEG
bukūn
possible
‘If I said I speak their language, it is not possible.’

Jonathan Owens

(10) Relative clause

balkallam le-əm be l-luqqa
 speak.IMPF.1SG to-3PL.M with DEF-language
 l-biyarifū-ha di
 REL-know.IMPF.3PL.M-3SG.F DEM.SG.F
 ‘I speak to them in the language which they know.’

2. Text referential, cataphoric.

da is used cataphorically to foreshadow a propositional expansion. In (11) the speaker is asked how he farms. Instead of answering directly, he introduces his answer with the cataphoric use of *da*, which is then expanded upon in the following independent proposition.

(11) a. kēf tihērit

how farm.IMPF.2SG.M
 ‘How do you farm?’

b. baharit da, al-hirāta l-wād-e tul di
 farm.IMPF.1SG DEM.SG.M DEF-farming DEF-one-F only DEM.SG.F
 d-duxun
 DEF-millet
 ‘How I farm? The one type of farming is only millet.’

3. Deictics.

A number of deictic words, mainly adverbs, are marked by demonstratives, in this case nearly always *da*. The deictics include *hassa* ‘now’, *dugut* ‘now’, *wakit/waqit* ‘now’, *tawwa* ‘previously, formerly’, *hine/hinēn* ‘here’, *awwal* ‘first, before’, *gabul* ‘previously, before’, *baʔad* ‘afterwards’, *alōm/alyōm* ‘today’, *amis* ‘yesterday’, *albāre* ‘yesterday evening’, *ambākir* ‘tomorrow’, *bukura* ‘day after tomorrow’, *məṇṇaṣabá* ‘in the morning’, *qādi* ‘there’, *hināk* ‘there’, *hajira* ‘(a place) away from here’, *bilhēn* ‘much’

(12) hajira da ma mašēt
 away DEM.SG.M NEG go.PRF.1SG
 ‘I didn’t go away anywhere.’

(13) albāre da as-sarārik daxalo
 yesterday DEM.SG.M DEF-thieves enter.PRF.3PL.M
 ‘Yesterday evening thieves broke in.’

4. Demonstratives mark pronouns, in this case often agreeing with the pronoun in terms of number and gender, and other demonstratives, where usually *da* occurs.

- (14) a. inti di ġibi le-i š-suqúl da
 2SG.F DEM.SG.F bring.IMP.SG.F to-OBL.1SG DEF-thing DEM.SG.M
 ‘You there bring me the watchamacallit.’
 a. ʔard gaydam dōla da kula ʔarab
 land Geidam DEM.PL.M DEM.SG.M also Arab
 ‘In the land around Geidam and the like are also Arabs.’

Basic attributes of these expanded functions can be given in cursory manner.

Concerning frequency, the occurrence of demonstratives in these functions on a token basis is high. For instance, there are 887 tokens of *qādi* ‘there’ in the corpus, of which 108 or 12% are marked by *da*. The highest percentages of demonstratives in these functions occur with dependent clauses and the 3sg pronouns *hu* ‘he’ and *hi* ‘she’. For *hu*, nearly 25% of all tokens occur with *da* (586/2407 24.3%). As far as the four innovative functions summarized above are concerned, a sample of 1318 tokens of *da* gathered from an arbitrary selection of 45 texts in the corpus reveals the data presented in Table 3. While the inherited referential functions constitute the largest single class, they make up only 53% of the total. The remaining 47% are functionally innovative.

Table 3: Functions of *da* in NA

Function	Percentage of total
Inherited functions	53.4
Entity referential: 557	42.3
Proposition-anaphoric: 146	11.1
Innovative functions	46.7
Cataphoric-propositional: 95	7.2
Dependent clause: 246	18.7
Adverbs/deictic: 158	12
Pronouns, demonstratives: 116	8.8

The syntactic, pragmatic and semantic nuances of using or not using the demonstratives in these innovative contexts have yet to be worked out. The two exam-

Jonathan Owens

ples in (15) and (16) illustrate different ways the innovative functions are integrated with other elements of the grammar.

Syntactically, for instance, based on the sample described above, *da* marks the end of about 30% of all conditional clauses. When it does not occur, its final clause boundary marking position commutes with an alternative pragmatically-marked element, such as the discourse marker *kula* ‘even’. (No tokens of **kula da* closing a conditional clause occur in the corpus).

- (15) kan qayyart-a kula
 if change.PRF.2SG.M-3SG.M DM
 ‘Even if you changed it.’

Pragmatically there are many instances where *da* has a focusing function, as in the following, where a mixed linguistic region ‘here’ is contrasted with another ‘there’, which is linguistically homogeneous.

- (16) nās gadé gadé kula hinēn katīrīn fi [qādi da] nafar-na
 people different different DM here many EXS there DEM.SG.M type-1PL
 nafara wāhid
 type one
 ‘Here there are a lot of different (types) but [over there] there is just our
 one ethnic group.’

The functions outlined in Table 3 are therefore both of high frequency and are systematically embedded in the syntax and pragmatics.

It should be intuitively clear that the functions in examples (9)–(16) are innovative in their systematicity relative to other varieties of Arabic. To show this in detail it would, however, be necessary to look at large-scale corpora of other Arabic dialects. This can very briefly be done with EA, which, as noted above, is an ancestral homeland of NA. The EA corpus is from *LDC Callhome* (Canavan et al. 1997), Nakano (1982), Behnstedt & Woidich (1987), and Woidich & Drop (2007), comprising about 417,000 words. It is thus of comparable size to the NA corpus. In this corpus there do occasionally occur collocations of pronoun + demonstrative in the same contexts as illustrated in (14), in particular as in (17).

- (17) hiwwa da lli mawgūd fandi-na
 3SG.M DEM.SG.M REL present at-1PL
 ‘That is what we have.’

It clearly, however, has a different functionality from NA pronoun + demonstrative. In EA the construction consistently is anaphoric to a previous proposition or situation, as in (17), where it introduces a previously-established topic to a following descriptive qualification. In 11 of the 58 tokens in the EA corpus it is followed by a relative clause, as in (17). Most tellingly, there are 2,677 *huwwa* (~ *hu*, *hū*, *hiwwa*, *hūwa*) tokens, of which only 58, or 2% are followed by *da* (~ *dah*, *dih*, *deh*, *dī*). This compares to the nearly 25% *hu* + *da* tokens in NA noted above. Moreover, in the NA sample, no tokens of *hu da* are followed by a relative clause. In this same statistical vein, the total number of singular proximal demonstratives in NA amounts to 16,774 tokens (14,591 *da*, 2,183 *dī*). In the EA corpus there are only 8,239 (4,996 *da*, 3,243 *dī*). Given that the corpora sizes are comparable (EA in fact a little larger), the demonstratives in NA are vastly over-proportional. This preponderance is due to *da*. Clearly there is a case to be answered for: what accounts for the vastly higher frequency of the 3sg.M demonstrative in NA? Recall, in answering this question, that behind the simple statistical comparison is a fundamental historical one as well. Ancestral NA came from ancestral EA. The initial populations, it needs to be assumed, had a demonstrative system like that of EA, and the majority of NA demonstrative tokens (see Table 3) still reflect this system. A blunt historical linguistic question is what caused the vast shift in frequencies.

From these initial, basic observations, it does not appear that the greatly expanded functionality of the demonstrative in NA can be explained by an increasing grammaticalization of the demonstrative.² This follows from two observations. First, the expanded functions of the demonstrative in Table 3 are, with the exception of the boundary-marking of dependent clauses (10), not those associated with the grammaticalization of demonstratives (e.g. the 17 trajectories of demonstratives in Diessel 1999). Secondly, NA and EA split over 400 years ago. One of the branches, represented by NA, underwent the considerable changes outlined here, whereas the other branch, EA, probably did not change at all (i.e. sentences such as (17) were probably present in EA in 1200, and before).³ There is thus no natural or inherent tendency for demonstratives to expand as in Table 3. It can thus be safely assumed that the expanded functionality of the NA

²I do not at all agree with Heine & Kuteva (2011) and Leddy-Cecere (this volume) that changes due to contact can be assimilated to a type of grammaticalization process, so the following contact-based account is independent of grammaticalization. Grammaticalization, in the original Meilletian sense, pertained only to internally-motivated changes.

³Cf. Damascus, which has an identical construction to that of EA. There are parallels in Classical Arabic as well, so this type of construction is probably proto-Arabic. If so, it only heightens the degree to which NA has innovated away from an original, stable structure.

Jonathan Owens

demonstrative was due to contact.

In fact, there is a large amount of *prima facie* evidence supporting this supposition. However, as is so frequently the case when one suspects pattern (metatypical)-type contact influence which is probably centuries old, support for the position will be indexical. Moreover, in the current case one is most probably dealing with a large-scale areal phenomenon in the Lake Chad area (and perhaps beyond) which encompasses well over 100 languages. In this summary chapter it will therefore have to suffice to rather peremptorily indicate that throughout the region there is a referential marker, sometimes a demonstrative, sometimes an article-like element, sometimes an element with both demonstrative and article-like properties, which consistently has the distribution of (9)–(16). Some languages have a better fit than others, and, of course, they will differ in detail in their language-internal functionality. A basic pattern is illustrated in (18) with Kanuri (Hutchison 1981: 47, 207, 218, 234, 241, 270), and summary references are made for Bagirmi, Wandala and Fali. So far as is known, Fali and Wandala had no significant contact with NA or its WSA relatives.

The Kanuri determinative *-da* has the following functions.

(18) Kanuri (Nilo-Saharan/Saharan)

Anaphoric entity reference

- a. obligatorily ends RC and optionally many adverbial clauses; = (9), (10)
- b. pronoun focus; = (14)
- c. marks adverbs; = (12), (13)

The only Kanuri structure missing from the list appears to be the propositional cataphoricity illustrated in (11).

Wandala (Frajzyngier 2012: 507–34, 603) has two morphemes: *-na* which is broadly glossed as a determiner and *-w* ‘that’. *-na*, besides marking entity reference, obligatorily marks the ends of a relative clause, and optionally a conditional (= (9), (10)); it occurs as an obligatory element in certain time/place adverbs (= (12), (13)); it is part of the previous mention marker *ɲán-na*; *ɲán* itself is said to originally be a third person singular pronoun, so there is a structural parallel to *hu + da*. *-w* functions as a topic marker that marks pronouns (= (14)).

In Fali (Niger-Congo/Adamawa) the demonstratives *gi/go* also obligatorily mark the end of relative and conditional clauses (= (9), (10)), subject focus (= (14)), and occur with some adverbs (= (12), (13)).

In Bagirmi a “determiner particle” *-na* is a constitutive part of the demonstrative *enna* < *et-na* ‘this’, and *-na* alone obligatorily marks the end of relative

clauses, and can emphasize pronouns, adverbs and entire sentences (Stevenson 1969: 40, 51, 54).

Areal features typically are not sensitive to language family, and this appears to be the case in this brief exemplification. Kanuri and Bagirmi are Nilo-Saharan, Wandala Chadic, Fali Niger-Congo and Arabic is Semitic. Only Wandala and Arabic are very distantly related genetically. Nonetheless, in all of the languages there is a deictic–referential marker (demonstrative, determiner, demonstrative–determiner) which, besides a classic deictic or anaphoric function, surfaces in an extended range of identical (cf. marking boundary of dependent clause) or similar (pronouns, adverbs)⁴ functions. These extended functions are precisely those which distinguish NA from other varieties of Arabic. The case for contact follows from two directions: in certain (not all) respects, NA deviates markedly from a putative ancestral source shared with EA, and where it does, its deviation corresponds broadly to analogous categories in co-territorial languages.

3.4 Semantics

The innovative distribution of the NA demonstratives is striking for the degree to which it appears to have raised the overall demonstrative token count, relative to EA. Discerning its presence in a text, however, is a straightforward matter. A much subtler, but no less pervasive instance of contact-based change pertains to idiomaticity. Like the demonstrative, this has a semantic and a formal aspect. Semantically, meanings emerge which are, for Arabic, unique, as in the following.

- (19) a. *rās al-bēt*
 head DEF-house
 ‘roof’
 b. *nādim rās-a*
 person head-3SG.M
 ‘an independent person, person of his own means’
- (20) a. *tallafo gaḷb-i*
 spoil.PR.F.3PL.M heart-1SG
 ‘They angered me.’

⁴The comparativist is limited to the extant reference grammars. These are in many instances excellent. Still, I suspect that they understate the flexibility of distribution of elements such as the deictic marker discussed here. *Mea culpa*, in Owens (1993: 88, 221, 235) the extended functions of the demonstrative described in this chapter for NA were treated in disparate sections, with no overall focus.

Jonathan Owens

- b. *gaḷb*-a *helu*
 heart-3SG.M sweet
 ‘He is happy.’

Formally the idioms are distinctive (as Arabic collocations) in bringing together lexemes which in other dialects would hardly co-occur, like [*tallaf* + *gaḷb*] or [*gaḷb* + *helu*]. The idiomatic meanings of the keywords (e.g. *tallaf*, *gaḷb*) are, in usage terms, often the typical usage for a given lexeme. In the NA corpus, for instance, of 101 tokens of *gaḷb* ‘heart’ all of them, 100%, are idiomatic. There is no reference to a physical heart. Similarly, *rās* is 80% idiomatic (247/308 tokens; Ritt-Benmimoun et al. 2017: 53). Thus, while idiomaticity has been consistently ignored as a theoretical issue in historical linguistics in general and in Arabic in particular, on a usage basis it is an integral aspect of understanding the lexical texture of the language.

Here as well NA is strikingly different from EA, as again can be determined from corpora-based comparison. In general, though both NA and EA share idiomatic keywords (*gaḷb*/*ʔalb* and *rās* are frequent in both for instance), their meanings and their collocational environments hardly overlap. For instance, in the EA corpus there are 110 tokens of *gaḷb*/*ʔalb* ‘heart’, of which 102 or 93% are idiomatic. This percentage closely parallels that of NA idiomatic *gaḷb*. The typical EA collocate of idiomatic *ʔalb*, however, is very different. The most frequent meaning is ‘center of X’, *ʔalb il-baḥr* ‘middle of the sea’. This meaning is entirely lacking in NA, and consequently collocates like *!gaḷb al-baḥar* (! = collocationally/semantically odd) are also lacking.

How different NA idiomaticity (meaning and collocational environment) is from EA was shown recently in Ritt-Benmimoun et al. (2017). There a three-way comparison was conducted between EA, southern Tunisian Arabic and NA, looking at three idiomatic keywords frequent in all three dialects: *rās*, *gaḷb* ‘heart’, and *ʕēn* ‘eye’. EA and southern Tunisian, though separated by a longer period of time (ca. 1035–present) than EA–NA (ca. 1300–present), showed a much higher identity of idiomatic structure than EA–NA (or NA–southern Tunisian). Both Egyptian (21a) and Tunisian Arabic (21b), for instance, maintain the same lexemes, same structure, same idiomaticity in a highly specific meaning.

- (21) a. Egyptian Arabic
 ḥaṭṭ rās-u fi t-turāb
 put.PRF.3SG.M head-3SG.M in DEF-ground
 ‘He humiliated him.’
 b. Tunisian Arabic

ḥaṭṭ-l-a rās-a fi t-trāb
 put.PROF.3SG.M-to-3SG.M head-3SG.M in DEF-ground
 ‘He humiliated him.’

These are nonsensical, or literal collocations in NA.

The comparison between EA and southern Tunisian Arabic serves as a similar baseline to comparing the overall demonstrative frequencies between EA and NA. The same question occurs. Why is NA different?

In this case the answer is even clearer than with the demonstrative. Essentially, NA has calqued its idiomatic structure (meaning and collocation) from Kanuri. The Kanuri of (19a) and (20b), for instance, are as in (22).

- (22) a. kəla fato-be
 head house-GEN
 ‘roof’
 b. kam kəla-nzə-ye
 person head-3SG.M-GEN
 ‘an independent person, person of his own means’

A ‘roof’ in both languages is the ‘head of a house’, an independent person is a ‘person of his head’, and so on, for something in the range of 70–80% of all the approximately 340 idioms studied (see Owens 1996; 2014; 2015; 2016b for details).

In summary, a large part of NA lexical structure is, as it were, not Arabic, but rather, as termed in Owens (1998), part of the Lake Chad idiomatic area. This identity, however, exists only at a semantic and collocational level. In their basic meaning, and their phonology, morphology and syntax, even in the context of idioms (Owens & Dodsworth 2017), the constituent lexemes *rās*, *bēt*, *tallaf*, *galb* etc. in NA are indistinguishable from any variety of Arabic at all.

There doubtless remains a good deal more systematic, contact-based correspondence between NA and languages of the Lake Chad area to be explored. The influence on NA is significant.

4 Conclusion

According to the historico-demographic background to NA, this variety did and does live with co-territorial languages, particularly Kanuri, today increasingly with Hausa, and in the past, Fulfulde and other smaller languages. NA bilingualism should, presumably, manifest itself in borrowing. Equally, NA speech com-

Jonathan Owens

munities have incorporated speakers of other languages into its fabric. The expectation here is that NA would be influenced via shift (imposition) from other languages.

In the domains summarized here, it is hard to discern a clear correlation between linguistic outcome and type of contact. There has been some phonological change, which in Van Coetsem's (1988; 2000) model is suggestive of change via shift (imposition), but the influence is limited to the features discussed in §3.1. What I believe is more striking than the contact-induced phonological change is the maintenance of inherited structures. NA still maintains a robust series of emphatics, has a non-reductive syllable structure reminiscent of, *inter alia*, Tihāma varieties, has classic distinguishing syllable structure attributes such as the *ga-hawa* syndrome (*ahamar* 'red') and the *bukura* syndrome (*bi-ǧiri* 'he runs'), to mention but a few. If the changes in (9)–(16) are due to imposition, it is equally clear that the “imposers” otherwise learned/learn a very normal Arabic.

Classic borrowing is moderate. The fact that discourse markers and conjunctions are token-wise frequent suggests that speakers were/are conversant in both Kanuri and Arabic. This does not, however, indicate whether these loans arose through imposers or borrowers. Moreover, to complicate matters even more, assuming Kanuri to have been the widespread *lingua franca* in the past, it would not need to have been native Kanuri speakers who imposed the Kanuri into Arabic. Speakers of Fulfulde, Kotoko, Malgwa or other languages would have been involved as well. As shown in Owens & Hassan (2010), discourse markers are prevalent in codeswitching, which here would be conducted by Arabs codeswitching between Arabic and Kanuri. From this scenario the discourse markers entered as borrowed elements.

The interpretation of demonstratives and idiomatic structure is equally ambiguous. The easiest development to envisage is L2 Arabic speakers imposing their L1 Kanuri, Fulfulde etc. usage onto their L2 Arabic. What makes this interpretation attractive is that it explains why in both cases such a massive importation of non-Arabic structure came into Arabic. As the name implies, these speakers could simply have imposed their own semantics and collocational alignment onto Arabic. Equally, however, it is not impossible that L1 Arabic speakers, fully bilingual in Kanuri and/or other languages simply shifted their Arabic usage to accommodate to their L2. Full fluency implies knowing idiomatic structure and the use of demonstratives, which the Arab borrowers could eventually incorporate into their own Arabic.

The only obvious common denominator to these musings is that the speakers would have been highly fluent in their respective L2s, whether L2 Arabic speak-

ers shifting to Arabic or L1 Arabic speakers fluent in Kanuri or other languages borrowing from their L2. The issue is only partly who the L1 and L2 speakers are. It is equally how well the populations knew/know Arabic/other languages, and how the high level of fluency produces the results shown.

Adding to the interpretive problem is that neither of the domains, idiomaticity or the expansion of demonstratives as it occurred in NA, have a comparative basis. Idiomaticity in the recent western linguistic tradition has been all but entirely subordinated to metaphor theory (Lakoff & Johnson 1999; see Haser 2005 for one critical perspective). It has received very little principled historical interpretation, and what work has been done (e.g. Sweetser 1990) tends to follow a Lakoffian paradigm and to be confined to European languages and to societies quite different from that of Nigerian Arabs. As far as demonstratives go, the little work that has been done on the languages co-territorial with NA (e.g. Kramer 2014: 141 on Fali), assume a grammaticalization of demonstrative usage *ab novo* via grammaticalization processes. Assuming such a perspective for the development of NA gives the lie to this simple assumption for the following reason. It would need to explain why the grammaticalization process did not take place in EA or other Arabic varieties, but did in NA, which is spoken in an area where the co-territorial languages, historically antecedent to Arabic, have the structures which NA acquired. If change via contact is the only plausible explanation for NA, it equally needs to be entertained for any language in the Lake Chad region.

Given so many open variables, it might be interesting to approach the issue from the opposite perspective, namely, what parts of language were not influenced by contact. Most of phonology was not, morphology hardly at all, syntax to a degree, basic vocabulary little.⁵ This minimally implies that if the contact changes were due to shift, the shifters in other domains (those where they did not impose idiomaticity or demonstrative usage) acquired a native-like competence in Arabic. In this respect it might be easier to envisage L1 Arabic borrowers maintaining these structures, and borrowing idiomaticity/demonstrative usage via their L2.

At the end of the day I think the range of questions evoked far surpasses the ability of currently-formulated linguistic theories of contact or language change, whether based on sociolinguistic or on cognitive perspectives (Lucas 2015: 523) to provide profound insight into how the obvious, and in some cases pervasive influence on NA via contact came about. It would be more fruitful to turn the question around and ask how rich data bases such as exist for NA, EA and some

⁵A Swadesh 100-word list gives something in the range of 79–83% cognacy with other varieties of Arabic.

Jonathan Owens

other Arabic dialects inform the overall issue of change via contact.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	N	noun
DEF	definite article	NA	Nigerian Arabic
DEM	demonstrative	NEG	negative
DM	discourse marker	NP	noun phrase
EA	Egyptian Arabic	OBL	oblique
EXS	existential	PL	plural
F	feminine	PRF	perfect
GEN	genitive	REL	relative
IMPF	imperfect	SG	singular
M	masculine	WSA	West Sudanic Arabic
N	number		

References

Behnstedt, Peter & Manfred Woidich. 1987. *Die Ägyptisch-arabischen Dialekte*. Vol. 3: *Delta Dialekte*. Wiesbaden: Reichert.

Braukämper, Ulrich. 1994. Notes on the origin of baggara arab culture with special reference to the shuwa. *Sprache und Geschichte in Afrika (SUGIA)* 14. 13–46.

Broß, Michael. 2007. L2 speakers, koineization and the spread of language norms: hausa in Maiduguri, nigeria. In Gudrun Mieke, Jonathan Owens & Manfred von Roncador (eds.), *Languages in African urban contexts*, 51–74. Hamburg: LIT.

Canavan, Alexandra, George Zipperlen & David Graff. 1997. *CALLHOME Egyptian Arabic Speech LDC97S45*. Web Download. Philadelphia, Linguistic Data Consortium.

Diessel, Holger. 1999. The morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology* 3. 1–46.

Frajzyngier, Zygmunt. 2012. *Wandala*. Berlin: Mouton de Gruyter.

Haser, Verena. 2005. *Metaphor, metonymy and experientialist philosophy: Challenging cognitive semantics*. Berlin: Mouton de Gruyter.

Heine, Bernd & Tania Kuteva. 2011. The areal dimension of grammaticalization. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of grammaticalization*, 291–301. Oxford: Oxford University Press.

- Hutchison, John. 1981. *The Kanuri language*. Madison: African Studies Program.
- Kramer, Raija. 2014. *Die Sprache der Fali in Nordkamerun*. Cologne: Rüdiger Köppe.
- Lakoff, George & Mark Johnson. 1999. *Philosophy in the flesh*. New York: Basic Books.
- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Manfredi, Stefano. 2010. *A grammatical description of kordofanian baggara Arabic*. Naples: Università degli Studi di Napoli ‘L’Orientale’ doctoral dissertation.
- Manfredi, Stefano. 2014. Demonstratives in a Bedouin arabic dialect of western sudan. *Folia Orientalia* 51. 27–50.
- Max Lock Group. 1976. *Max Lock Group report: Maiduguri: Surveys and planning reports for Borno, Bauchi and Gongola State governments*. London: Westminster.
- Nakano, Aki’o. 1982. *Folktales of Lower Egypt: Texts in Egyptian Arabic*. Tokyo: Institute for the Study of Languages, Cultures of Asia & Africa.
- Owens, Jonathan. Forthcoming. Arabic. In Salikoko S. Mufwene & Anna María Escobar (eds.), *The Cambridge handbook of language contact*. Cambridge: Cambridge University Press.
- Owens, Jonathan. 1993. *A grammar of Nigerian Arabic*. Wiesbaden: Harrassowitz.
- Owens, Jonathan (ed.). 1994a. *Arabs and Arabic in the Lake Chad region*. Special issue of *Sprache und Geschichte in Afrika (SUGIA)* 14.
- Owens, Jonathan. 1994b. Nigerian arabic in comparative perspective. *Sprache und Geschichte in Afrika (SUGIA)* 14. 85–176.
- Owens, Jonathan. 1996. Idiomatic structure and the theory of genetic relationship. *Diachronica* 13(2). 283–318.
- Owens, Jonathan. 1998. *Neighborhood and ancestry: Variation in the spoken Arabic of Maiduguri (Nigeria)*. Amsterdam: John Benjamins.
- Owens, Jonathan. 2000. Loanwords in Nigerian arabic: a quantitative approach. In Jonathan Owens (ed.), *Arabic as a minority language*, 259–346. Berlin: De Gruyter.
- Owens, Jonathan. 2002. Processing the world piece by piece: iconicity, lexical insertion and possessives in Nigerian arabic codeswitching. *Language Variation and Change* 14(2). 173–209.
- Owens, Jonathan. 2006. *A linguistic history of Arabic*. Oxford: Oxford University Press.
- Owens, Jonathan. 2013. The historical linguistics of the intrusive -n in Arabic and west semitic. *Journal of the American Oriental Society* 133(2). 217–248.

Jonathan Owens

- Owens, Jonathan. 2014. Many heads are better than one: the spread of motivated opacity via contact. *Linguistics* 52(1). 125–165.
- Owens, Jonathan. 2015. Idioms, polysemy, context: a model based on Nigerian arabic. *Anthropological Linguistics* 57(1). 46–98.
- Owens, Jonathan. 2016a. Dia-planar diffusion: reconstructing early aramaic–Arabic language contact. In Manuel Sartori, Manuela Giolfo & Philippe Casuto (eds.), *Approaches to the history and dialectology of Arabic in honor of Pierre Larcher*, 77–101. Leiden: Brill.
- Owens, Jonathan. 2016b. The lexical nature of idioms. *Language Sciences* 57. 49–69.
- Owens, Jonathan & Robin Dodsworth. 2017. Semantic mapping: what happens to idioms in discourse. *Linguistics* 55(3). 641–682.
- Owens, Jonathan & Jidda Hassan. 2004. Remarks on ideophones in Nigerian arabic. In Martine Haak, Rudolf de Jong & Kees Versteegh (eds.), *A Festschrift for Manfred Woidich*, 207–220. Leiden: Brill.
- Owens, Jonathan & Jidda Hassan. 2010. Conversation markers in Arabic–hausa codeswitching: saliency and language hierarchies. In Jonathan Owens & Alaa Elgibali (eds.), *Information structure in spoken Arabic*, 207–242. London: Routledge.
- Owens, Jonathan & Jidda Hassan. 2011–17. *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/Idiomacity_lexical_realignment_and_semantic_change_in_spoken_arabic/Nigerian_Arabic/index.html.
- Retsö, Jan. 2013. What is arabic? In Jonathan Owens (ed.), *The Oxford handbook of Arabic linguistics*, 433–450. Oxford: Oxford University Press.
- Ritt-Benmimoun, Veronika, Smaranda Grigore, Jocelyne Owens & Jonathan Owens. 2017. Three idioms, three dialects, one history: egyptian, nigerian and tunisian arabic. In Veronika Ritt-Benmimoun (ed.), *Tunisian and libyan arabic dialects: common trends – recent developments – diachronic aspects*, 43–84. Zaragoza: Prensas de la Universidad de Zaragoza.
- Stevenson, R. C. 1969. *Bagirmi grammar*. Khartoum: Sudan Research Unit, University of Khartoum.
- Sweetser, Eve. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

7 *Nigerian Arabic*

- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Woidich, Manfred & Hanke Drop. 2007. *ilBaḥariyya – Grammatik und Texte*. Wiesbaden: Harrassowitz.

Chapter 8

Arabic pidgins and creoles

Andrei Avram

University of Bucharest

The chapter is an overview of eight Arabic-lexifier pidgins and creoles: Turku, Bon-gor Arabic, Juba Arabic, Kinubi, Pidgin Madame, Jordanian Pidgin Arabic, Roman-ian Pidgin Arabic, and Gulf Pidgin Arabic. The examples illustrate a number of selected features of these varieties. The focus is on two types of transfer, imposi-tion and borrowing, within the framework outlined by Van Coetsem (1988; 2000; 2003) and Winford (2005; 2008).

1 Introduction

This chapter aims to illustrate the emergence of Arabic-lexifier pidgins and cre-oles for which the contact situation – i.e. socio-historical context, the agents of change, and the languages involved – is at least relatively well known.

The varieties considered can be classified into two groups, in geographical, historical and developmental terms: the Sudanic pidgins and creoles, and the im-migrant pidgins in various Arab countries. Geographically, the Sudanic varieties developed in Africa – in present-day South Sudan, Chad, Uganda, and Kenya. His-torically, the Sudanic varieties derive from a putative common ancestor, a pidgin that emerged in southern Sudan, in the first half of the nineteenth century. Vari-ous Turkish–Egyptian military expeditions between 1820 and 1840 opened south-ern Sudan for the slave trade. Permanent camps were set up soon after by slave traders in the White Nile Basin, Bahr el-Ghazal and Equatoria Province, inhab-ited by an Arabic-speaking minority and a huge majority of slaves from various ethnic and linguistic backgrounds. After 1850, the slave traders’ settlements were turned into military camps in which a military pidgin emerged, which is tradi-tionally referred to as “Common Sudanic Pidgin Creole Arabic” (Tosco & Man-fredi 2013: 253). Two subgroups of Sudanic varieties are recognized: the western



Andrei Avram

branch, consisting of Turku and Bongor Arabic (in Chad), and the eastern one, made up of Juba Arabic (in Sudan) and Kinubi (spoken in Uganda and Kenya).

Immigrant pidgins emerged in the eastern part of the Arab World, in Lebanon, Jordan, Iraq and the countries of the Arab Gulf. Historically, these do not go back more than 50 years. All these varieties are incipient pidgins.

The contact situations illustrated presuppose: (i) a SL and a RL; (ii) agents of contact-induced change, who may be either SL or RL speakers; (iii) a psycholinguistically dominant language, which is not necessarily a socially dominant language (Van Coetsem 1988; 1995; 2000; 2003; Winford 2005; 2008). A distinction is made between two types of transfer: imposition and borrowing (Van Coetsem 1988; 2000; 2003). Imposition involves SL-dominant speakers as agents (SL agentivity), is typical of second language acquisition, and induces changes mostly in phonology and syntax, although it may also include transfer of lexical items from the dominant SL into the non-dominant RL (Van Coetsem 1995: 18; Winford 2005: 376). Borrowing normally involves RL-dominant speakers as agents (RL agentivity), it typically targets lexical items, but may also include transfer of morphological material from a non-dominant SL into the dominant RL.

In light of their sociolinguistic history, the varieties considered all emerged under conditions of untutored, short term, second language acquisition by adults dominant in their socially subordinate SLs. Second language acquisition, *a fortiori* with adults, triggers processes such as imposition via SL agentivity (i.e. substrate influence), simplification (Trudgill 2011: 40, 101) – also known as restructuring (Lucas 2015: 529) – as well as language-internal (i.e. non-contact-induced) developments such as grammatical reanalysis (Winford 2005: 415).

As in Manfredi (2018), the focus of this chapter is on imposition and borrowing. It does not illustrate restructuring which does not involve any kind of transfer, but often involves a reduction in complexity (Lucas 2015: 529). In the case of Arabic pidgins and creoles, restructuring is manifest in the domain of morphology, in, for example, the loss of the Arabic verbal affixes and of the nominal and verbal derivation strategies (Miller 1993).

The examples are illustrative only of selected contact-induced features of Arabic pidgins and creoles and their number has been kept to a reasonable minimum. The examples from Arabic and the pidgins and creoles considered appear in a uniform system of transliteration.

The chapter is organized as follows. §2 and §3 are concerned with Sudanic pidgins and creoles. §4, on the other hand, deals with Arabic immigrant varieties. §5 summarizes the findings and introduces issues for further research.

2 Turku and Bongor Arabic

2.1 Current state and historical development

Turku is an extinct pidgin, formerly spoken in the Chari–Bagirmi region in western Chad (Muraz 1926). After the abolition of slavery by the Turkish–Egyptian government in 1879, the Nile Nubian trader Rabeh withdrew with his slave soldiers into Chad. From a sociolinguistic point of view, Turku was initially a military pidgin. However, it later became one of three trade languages in what was then French Equatorial Africa, along with Sango and Bangala (Tosco & Owens 1993: 183). Turku was a stable pidgin which does not appear to have creolized (Tosco & Owens 1993).

Bongor Arabic is spoken in southwestern Chad, in and around the town of Bongor, the capital of the Mayo–Kebbi Est region, close to the border with Cameroon (Luffin 2013). Given the many structural features it shares with Turku, it is plausible to assume that Bongor Arabic developed from the former. Sociolinguistically, Bongor Arabic is a trade pidgin, used by the local Masa and Tupuri populations with Arabic-speaking traders. It is currently a stable pidgin, but it exhibits features indicative of depidginization under the influence of Chadian Arabic. No information about the number of speakers is available.

2.2 Contact languages

The lexifier language of Turku and Bongor Arabic is Western Sudanic Arabic. The substratal input was provided by languages of various genetic affiliations: Nilo-Saharan – e.g. Bagirmi, Mbay, Ngambay, Sar, Sara (Central Sudanic), Kanuri (Western Saharan); Afro-Asiatic – Hausa (West Chadic); Niger-Congo – Fulfulde. In the case of Turku, an additional contributor was the creole language Sango. Both in Turku and in Bongor Arabic there is also adstratal input from French. The adstrate of Bongor Arabic additionally includes two languages: Masa (Nilo-Saharan, Western Chadic) and Tupuri (Niger-Congo).

2.3 Contact-induced changes

2.3.1 Phonology

The substrate languages do not have /ħ/, which is generally replaced by [k]: Turku *kamsa* ‘five’ < ChA *ḥamsa*; Bongor Arabic *kídma* ‘work’ < ChA *ḥidma*. Many of the substrate languages do not have /f/, which is substituted with [p]

Andrei Avram

or perhaps [ɸ],¹ e.g. Turku *pfil* ‘elephant’ < ChA *fil*. In French loanwords, the reflexes of /v/ are either [b] or [w]: Bongor Arabic *boté* ‘to vote’ < French *voter*, *wotír* ‘car’ < French *voiture*.

The consonants [ɲ] and [ŋ] occur only in loanwords: Turku *konpanye* ‘company’ < French *compagnie*, *ngari* ‘manioc’ < Mbay *ngàrì*; Bongor Arabic *ngambáy* ‘Ngambay’ < Ngambay *ngàmbáy*; [v] and [ʒ] occur only in phonologically non-integrated words of French origin: Turku *sivil* ‘civilian’ < French *civil*; Bongor Arabic *žurnalíst* ‘journalist’ < French *journaliste*.

Variation affects several consonants. For instance, [f] occurs in variation with [b] or [p]: Turku *fīšan* ~ *bīšan* ‘because’; Bongor Arabic *māfi* ~ *māpi* ‘NEG’ < ChA *mā fī*, *sofēr* ~ *sopér* ‘driver’ < French *chauffeur*. Most of the substrate languages do not have /š/, which accounts for [ʃ] ~ [s] variation, in words with either etymological /s/ or /ʃ/: Turku *gasi* ~ *gaši* ‘expensive’ < ChA *gāsī*, *biriš* ~ *biris* ‘mat’ < ChA *birīš*; Bongor Arabic *māši* ~ *māsi* ‘go’. The usual reflexes of French /ʒ/, absent from the phonological inventories of the substrate languages, are [z], [dʒ] and [s] respectively: Turku *ģinenal* ‘general’ < French *général*, *suska* ‘until’ < French *jusqu’à*; Bongor Arabic *zúska* ‘when, during’ < French *jusqu’à* ‘until’.

Finally, vowel length is not distinctive: Turku, Bongor Arabic *kalam* ‘speech; speak’ < ChA *kalām* ‘speech’.

2.3.2 Morphology

On current evidence (Luffin 2013: 180–181), Bongor Arabic exhibits signs of depidginization under the influence of Chadian Arabic. The most striking instance of this is the use of pronominal suffixes, unique among Arabic-lexifier pidgins and creoles:

- (1) Bongor Arabic (Luffin 2013: 180)
 índi gáy árifu úsum-i
 2SG IMPF know name-POSS.1SG
 ‘You know my name.’

Also, verbal affixes are sporadically used:

- (2) Bongor Arabic (Luffin 2013: 181)

¹Given the transcription with <pf> in Muraz (1926: 168).

8 Arabic pidgins and creoles

- a. ána ma n-árfa
1SG NEG 1SG-know
'I don't know.'
- b. anína rikíb-na wotír da sáwa
1PL ride.PRF-1PL car PROX together
'We took the car together.'

These cases might be analyzed as borrowing under *sui generis* RL agentivity, whereby morphological material from a non-dominant SL is imported into a non-dominant (second) RL.

2.3.3 Lexicon

A part of the non-Arabic vocabulary of Turku can be traced back to its substrate languages (Avram 2019). Most of the loanwords are from Sara-Bagirmi languages: *adinbang* 'eunuch' < Bagirmi *ádim mbàṅ* 'servant of the sultan'; *gao* 'hunter' < Sar *gáw*; *ngari* 'manioc' < Mbay *ngàrì*. The second most significant important contributor is Sango: *kay* 'paddle' < Sango *kâi*, *tipoy* 'carrying hammock' < *típóí*. A few words can be traced to Fulfulde and Kanuri: *kelkelbanği* 'golden beads' < Fulfulde *kelkel-banja*; *wélik* 'lightning' < Kanuri *wulak* 'flash of lightning'. In a number of cases, the exact SL cannot be established: *koporo* '0.10 Francs' < Fulfulde, Sango, Sara *koporo* 'coin'; *gurumba* 'hat' < Hausa *gurúmba*, Kanuri *gurumbá*. As for Bongor Arabic, its African adstrate languages have contributed only a few loanwords, such as *bursdíya* 'Monday'. There are also loanwords from French. In Turku most of these relate to the military (Tosco & Owens 1993: 262–263), e.g. Turku *itenan* 'lieutenant' < French *lieutenant*, *permişon* 'permission' < French *permission*. In addition to nouns, French loanwords include some verbs, such as Bongor Arabic *komandé* 'order' < French *commander*, and at least one function word, Turku *suska*, Bongor Arabic *zúska* 'when, during' < French *jusqu'à* 'until'.

The substratal influence on Turku can also be seen in a number of compound calques (Avram 2019; Manfredi, this volume). Some of these are modelled on Sara-Bagirmi languages: *bahr gum* 'rising water', cf. Ngambay *màn-kàw*, lit. 'river goes'; *nugra ana asal* 'beehive', cf. Ngambay *bòlè-tànji*, lit. 'hole (of) honey'. Other calques have equivalents in several SLs, such as *nugra hağer* 'cave', lit. 'hole mountain/stone', cf. Kanuri *kûl kau-be* lit. 'cavity mountain-of', Ngambay *bòlò-mbàl* lit. 'hole mountain', Sango *dûtêně* lit. 'hole stone'.

Andrei Avram

3 Juba Arabic and Kinubi

3.1 Current state and historical development

Juba Arabic is mainly spoken in South Sudan; there are also diaspora communities, mostly in Sudan and Egypt. Two main reasons make it difficult to estimate its number of speakers. Firstly, while Juba Arabic is spoken as a primary language by 47% of the population of Juba, the capital city of South Sudan, it is also used as a second or third language by the majority of the population of the country (Manfredi 2017: 7). Secondly, the long coexistence of Juba Arabic with Sudanese Arabic, its main lexifier language, has led to the emergence of a continuum ranging from basilectal, through mesolectal, to acrolectal varieties; delimiting acrolectal Juba Arabic from Arabic is no easy task, particularly in the case of the large diaspora communities in Khartoum and Cairo.

Juba Arabic emerged as a military pidgin. Sociolinguistically, it is today an inclusive identity marker for the ethnically and linguistically diverse population of South Sudan (Tosco & Manfredi 2013: 507). Developmentally, Juba Arabic is a pidgincreole.²

The Mahdist revolt, which started in 1881, eventually brought to an end Turkish–Egyptian control over Equatoria, in southern Sudan. Following an invasion by Mahdist rebels, the governor fled to Uganda, accompanied by slave soldiers loyal to the central government. These soldiers subsequently became the backbone of the British King’s African Rifles. While some of the troops remained in Uganda, others were moved to Kenya and Tanzania. This led to the dialectal division between Ugandan and Kenyan Kinubi. Like Juba Arabic, therefore, Kinubi started out as a military pidgin, then underwent stabilization and expansion. Today, however, Kinubi is the only Arabic-lexifier fully creolized variety, that is, a native language for its entire speech community.

Kinubi is spoken in Uganda and in Kenya. The number of speakers of Kinubi is a matter of debate. Ugandan Kinubi was spoken by some 15,000 people, according to the 1991 census, and Kenyan Kinubi by an estimated 10,000 in 2005. However, other estimates put the combined number of speakers at about 50,000. The largest communities of Kinubi speakers are in Bombo (Uganda), Nairobi (the Kibera neighbourhood) and Mombasa (Kenya).

² A pidgincreole is “a former pidgin that has become the main language of a speech community and/or a mother tongue for some of its speakers” (Bakker 2008: 131).

3.2 Contact languages

The main lexifier language of Juba Arabic is Sudanese Arabic, with some input from Egyptian Arabic and Western Sudanic dialects as well. The substrate is represented by a relatively large number of languages, belonging to super-phylums, Nilo-Saharan and Niger-Congo. The former includes Eastern Sudanic languages, such as Bari, Lotuho (Eastern Nilotic), Acholi, Belanda Bor, Dinka, Jur, Nuer, Pări, Shilluk (Western Nilotic), Didinga (Surmic), and Central Sudanic languages, such as Avokaya, Baka, Bongo, Ma'di, Moru; the Niger-Congo super-phylum is represented by, for example, Zande and Mundu. The main substrate language is considered to be Bari, including its dialects Kakwa, Kuku, Pojulu, and Mundari.³

Given its sociolinguistic history, Kinubi shares much of its substrate with Juba Arabic. However, the substrate of Ugandan Kinubi additionally includes Eastern Sudanic languages, such as Alur, Luo (Western Nilotic), and Central Sudanic languages such as Mamvu, Lendu and Lugbara (Owens 1997: 161; Wellens 2003: 207), spoken in Uganda. Unlike Juba Arabic, Kinubi also exhibits the effects of the adstratal influence exerted by two Bantu languages, Luganda – particularly in Ugandan Kinubi – and Swahili – particularly in Kenyan Kinubi. One other language that should be mentioned is English, official both in Uganda and in Kenya.

3.3 Contact-induced changes

3.3.1 Phonology

A number of consonants found in Arabic, but absent from the phonological inventories of the substrate languages are either deleted or substituted. Consider the reflexes of pharyngeals: *háfla* 'feast' < SA *ḥafla*; *árabi* 'Arabic' < SA *ʔarabī*. The pharyngealized consonants are replaced by their plain counterparts: *towíl* 'long' < SA *ṭawīl*; *dul* 'shadow' < SA *ḍull*; *súlba* 'hip' < SA *ṣulba*; *zúlum* 'to anger' < SA *ẓulum*. The velar fricatives of Arabic are always replaced by velar stops: *kábara* 'piece of news' < SA *ḥabar*; *šókol* 'work' < SA *šoyol*, *gárib* 'west' < SA *ʔar(i)b*.

As in Juba Arabic, the pharyngeals of Arabic are either replaced or lost in Kinubi (Owens 1985: 10; Wellens 2003: 209–212). The earliest records of Ugandan Kinubi⁴ are replete with illustrative examples (Avram 2017a): *hağa* 'thing' < SA *ḥāğa*, *aram* 'thief' < SA *ḥarāmi*, *lib* < 'to play' < SA *liʔib*. The pharyngealized consonants are replaced by their plain counterparts, as in these examples from

³Sometimes considered to be separate languages (Wellens 2003: 207).

⁴The main ones are: Cook (1905), Jenkins (1909), Meldon (1913), and Owen & Keane (1915).

Andrei Avram

early Ugandan Kinubi: *towil* ‘long’ < SA *ṭawīl*; *dulu* ‘shadow’ < SA *ḍull*, *hisiba* ‘measles’ < SA *ḥiṣba*; *zūlm* ‘to anger’ < SA *ẓulum*. Like Juba Arabic, Kinubi substitutes velar stops for the Arabic velar fricatives. Consider the following early Ugandan Kinubi forms: *kidima* ‘work’ < SA *ḥidma*; *šokolo* ‘work’ < SA *šoyol*, *balago* ‘commandment’ < SA *balāy* ‘message’. Substratal influence also accounts for consonant degemination, given that the substrate languages “lack these in all but a few morphonologically determined contexts” (Owens 1997: 162).

Substratal influence can also be seen in the occurrence of certain consonants. Consider first /b/ and /d/: Juba Arabic *d’éngele* ‘liver’ < Bari *denggele*; Juba Arabic *b’ónḡo* ‘pumpkin’ < Bongo *b’onḡo*. The other consonants which occur only in loanwords from the substrate and/or adstrate languages are [p] [v], [tʃ], [ɲ], and [ŋ]: Kinubi *lípa* ‘to pay’ < Swahili *-lipa*; Kinubi *camp* ‘camp’ < English *camp*; Kinubi *víta* ‘war’ < Swahili *vita*; Juba Arabic *čam* ‘food’ < Acholi, Belanda Bor, Jur *čama*, Juba Arabic *čayniz* < English *Chinese*, Kinubi *čay* ‘tea’ < Swahili *chai*; Juba Arabic *nyékem*, Kinubi *nyékem* ‘chin’ < Bari *nyékem*, Kinubi *nyánya* ‘tomato’ < Swahili *nyanya*; Juba Arabic *ḡun* ‘divinity’ < Bari *ngun*. The integration of these phonemes is thus a result of borrowing (under RL agentivity) rather than of imposition.

The following instances of consonant variation are more common in Juba Arabic (Manfredi 2017: 25–27). The most frequent is [ʃ] ~ [s]: *geš* ~ *ges* ‘grass’. Further, [z] is in variation with [dʒ] before /o/ and /a/: *zówḡu* ~ *ḡówḡu* ‘to marry’, *záman* ~ *ḡáman* ‘time; when’. There is also [p] ~ [f] variation in word-initial position, including in loanwords: *poḡúlu* ~ *foḡúlu* ‘Pojulu’, *prótestan* ~ *frótestan* ‘Protestant’. Finally, the phoneme /f/ may also be phonetically realized as [p]: *nédifu* ~ *nédipu* ‘to clean’. Of these cases of variation, the latter has been specifically attributed to substratal influence from Bari (Miller 1989; Manfredi 2017). It might be argued, however, that all these instances of consonant variation reflect the influence of the substrate languages, regardless of their genetic affiliations. The following do not have /š/: Acholi, Avokaya, Baka, Bari, Belanda Bor, Bongo, Dinka, Jur, Lotuho, Ma’di, Moru, Mundu, Nuer, Pări, Shilluk, Zande. Of these, Acholi, Belanda Bor, Bongo, Dinka, Jur, Nuer, Pări and Shilluk do not have /s/ either. A number of substrate languages do not have /z/: Acholi, Bongo, Belanda Bor, Dinka, Jur, Lotuho, Nuer, Pări, and Shilluk. All of these, however, have /dʒ/. Finally, /f/ is not part of the phonological inventory of Acholi, Bongo, Dinka, Jur, Nuer, Pări, and Shilluk, which do, however, have /p/. Given the intricacies of the distribution of /š/, /s/, /z/, /dʒ/, /f/, and /p/ across the substrate languages, the types of variation illustrated are not surprising.

As in Juba Arabic, [ʃ] is in variation with [s] in Kinubi (Owens 1985: 237;

8 Arabic pidgins and creoles

Owens 1997: 161; Wellens 2003: 38; Luffin 2005: 62; Avram 2017a): early Ugandan Kinubi *šabaka* ~ *sabaka* ‘net’. Although it is etymological /š/ which is typically subject to variation, occasionally this also applies to etymological /s/: early Ugandan Kinubi *sikin* ~ *šekin* ‘knife’ < SA *sikkīn* (Avram 2017a) and modern Kenyan Kinubi *fluš* ~ *flus* ‘money’ < SA *fulūs* (Luffin 2005: 63). Note that [š] ~ [s] variation also extends to loanwords from Swahili (Wellens 2003: 80; Luffin 2005: 63; Avram 2017a): early Ugandan Kinubi *šamba* ~ *samba* ‘field’ < Swahili *shamba*. Like Juba Arabic, Kinubi exhibits [z] ~ [ɟ] variation (Owens 1985: 235; Owens 1997: 161; Wellens 2003: 215; Luffin 2005: 63; Avram 2017a): early Ugandan Kinubi *ǧalan* ~ *zalan* ‘angry’ < SA *zaʕlān*. However, unlike Juba Arabic, in Kinubi the [z] ~ [ɟ] variation also occurs before the two front vowels /i/ and /e/: *ze* ~ *ǧe* ‘as’, early Ugandan Kinubi *aṅil* ~ *enzil* ‘descend’. According to Owens (1997: 161), this “is due perhaps to Bari substratal influence, since Bari has only *j*, not *z*.” In fact, as in the case of Juba Arabic, the same is true of several other substrate languages. Finally, there are instances of [l] ~ [r] variation (Wellens 2003: 214; Luffin 2005: 65), affecting both etymological /l/ and etymological /r/ in Arabic-derived words, e.g. *tále* ~ *táre* ‘go out’, *gerí* ~ *gelí* ‘near’, and in borrowings, e.g. Ugandan Kinubi *čálo* ~ *čáro* ‘village’ < Luganda *e-kyalo*; Kenyan Kinubi *tumbíli* ~ *tumbíri* ‘monkey’ < Swahili *tumbili*. This variation seems to reflect the influence of Luganda and Swahili. In the former, [l] and [r] are in complementary distribution, with [r] occurring after the front vowels /i/ and /e/, and [l] elsewhere (Wellens 2003: 214), while in the latter [l] and [r] are in free variation (Luffin 2014: 79).

As in the substrate languages, there is no distinction between short and long vowels: Juba Arabic *sudáni* ‘Sudanese’ < SA *sudānī*, Kinubi *kabír* ‘big’ < SA *kabīr*.

3.3.2 Morphology

Apart from the Arabic-derived plural suffixes *-at* and *-in*, Juba Arabic uses the plural marker of Bari origin *-ǧín* (Nakao 2012: 131; Manfredi 2014a: 58), which is attached only to loanwords from local languages:

(3) Juba Arabic (Manfredi 2014a: 58)

- a. *kəɾɔɔ-ǧín* (< Bari *kəɾɔɔ*)
leaf-PL
‘leaves’
- b. *beng-ǧín* (< Dinka *beng*)
chief-PL
‘chiefs’

Andrei Avram

- c. b'angiri-ǵín (< Zande *b'angiri*)
cheek-PL
'cheeks'

Another phenomenon worth mentioning is the occurrence in the speech of young urban speakers of hybrid forms, which consist of the Bari relativizer *lo-* and a noun either from Arabic or from one of the substrate/adstrate languages (Nakao 2012: 131). Note, however, that there is a functional overlap between Bari *lo-* and Sudanese Arabic *abu*.

- (4) Juba Arabic (Manfredi 2017: 46)
 - a. lo-beléde (< Bari *lo-* + SA *beled*)
REL-country
'peasant'
 - b. lo-pómbe (< Bari *lo-* + Swahili *pombe*)
REL-alcohol
'drunkard'

Given that a relatively large number of Bari-derived words contain *lo-* (Miller 1989; Manfredi 2017: 46), the examples in (4) confirm the fact that morphological innovations are typically introduced through lexical borrowings via RL agentivity, and subsequently become productive in the RL.

Note, finally, that most of the speakers who use the plural marker *-ǵín* and the relativizer *lo-* are dominant in Juba Arabic. These cases therefore confirm the fact that RL monolinguals can be agents of borrowing (Van Coetsem 1988: 10).

A small number of Kinubi nouns borrowed from Swahili exhibit the Bantu nominal classifiers:

- (5) Kinubi (Wellens 2003: 57)
 - a. **mu-zé wa-zé**
NC1-old man NC2-old man
'old man, old men'
 - b. **mu-zukú wa-zukú**
NC1-grandchild NC2-grandchild
'grandchild, grandchildren'
 - c. **m-zúngu wa-zúngu**
NC1-European NC2-European
'European, Europeans'

3.3.3 Syntax

Bureng Vincent (1986: 77) first noted the similarity between the prototypical passive construction in Juba Arabic and its Bari counterpart:

(6) Juba Arabic (Bureng Vincent 1986: 77)

- a. bágara áyinu **ma** Wáni
 cow see.PAST with Wani
 ‘The cow has been seen by Wani.’
- b. Bari (Bureng Vincent 1986: 77)
 kítʒɪ a metà kò Wàni
 cow PAST see with Wani
 ‘The cow has been seen by Wani.’

As can be seen, in both Juba Arabic and Bari the agent is introduced by the comitative preposition ‘with’. This is a case of lexico-syntactic imposition via identification of SL and RL lexemes (Manfredi 2018: 415): the Juba Arabic lexical entry *ma* is derived from Sudanese Arabic *maʃ*, but its semantics reflects the influence of Bari *kò*. The same is true of Kenyan Kinubi:

(7) Kinubi (Luffin 2005: 230)

- yal-á al akulú **ma** nas tomsá
 child-PL REL eat.PAST.PASS with PL crocodile
 ‘The children who were eaten by a crocodile.’

Consider next the syntax of numerals in Kinubi (Wellens 2003: 90; Luffin 2014: 309). Their post-nominal placement is calqued on Swahili:

(8) Kinubi (Luffin 2014: 309)

- wéle kámsa ma baná árba
 boy five with girl.PL four
 ‘Five boys and four girls.’

(9) Swahili (Luffin 2014: 309)

- miti mia tatu
 tree hundred three
 ‘Three hundred trees.’

With cardinal numerals, the order is hundred + unit and thousand + unit respectively:

Andrei Avram

- (10) Kinubi (Luffin 2014: 309)
 elf wáy
 thousand one
 ‘one thousand’

Kinubi thus follows the Swahili model:

- (11) Swahili (Luffin 2014: 309)
 elfu moja
 thousand one
 ‘one thousand’

Consider also a case of syntactic change induced by lexical calquing. Juba Arabic (*fu*)*wata* ‘ground’ functions as an impersonal subject in weather expressions:

- (12) Juba Arabic (Nakao 2012: 141)
 (fu)watá súkun
 ground hot
 ‘It is hot.’

Nakao (2012: 141) shows that this is also the case in Acholi and Ma’di:

- (13) Acholi (Nakao 2012: 141)
 piiny lyeeet
 ground warm
 ‘It is warm.’

- (14) Ma’di (Nakao 2012: 141)
 vu aci
 ground hot
 ‘It is hot.’

In fact, these types of sentences are widespread in Western Nilotic substrate languages, such as Dinka, Jur, Pări, and Shilluk:

- (15) Dinka (Nebel 1979: 202)
 piny a-tuc
 ground 3SG-warm
 ‘It is warm.’

8 Arabic pidgins and creoles

In both Juba Arabic and Kinubi *ras* ‘head’ also occurs in the complex preposition *fi ras* ‘on’:

- (16) a. Juba Arabic (Nakao 2012: 141)
 merísa fī fī ras terebéza
 beer EXS on head table
 ‘The beer is on the table.’
- b. Kinubi (Wellens 2003: 159)
 fī rás séder
 on head tree
 ‘On top of the tree.’

Nakao (2012: 141) attributes this function of *ras* to substratal influence from Acholi and Ma’di:

- (17) Acholi (Nakao 2012: 141)
 cib wi-meja
 put head-table
 ‘Put it on the table!’

However, other possible sources include Western Nilotic languages such as Belanda Bor, Jur, Pāri and Shilluk:

- (18) Jur (Pozzati & Panza 1993: 342)
 kedh ŋo wi tarabesa
 put 3SG head table
 ‘Put it on the table.’

Moreover, a preposition ‘on’ derived from the noun ‘head’ is also attested in Bongo (Central Sudanic) and Zande (Niger-Congo):

- (19) Bongo (Moi et al. 2014: 39)
 ba do mbaa
 3SG on car
 ‘He is on a car.’
- (20) Zande (De Angelis 2002: 288)
 mo mai he ri ngua
 2SG put 3SG on wood
 ‘Put it on the wood.’

Andrei Avram

The verb *gal/gale/gali* ‘say’ is used in Juba Arabic and Ugandan Nubi as a complementizer, with *verba dicendi* and verbs of cognition:

- (21) a. Juba Arabic (Miller 2001: 469)
 úwo kélem gal úwo bi-ǵa
 3SG speak COMP 3SG IRR-come
 ‘He said that he would come.’
 b. Ugandan Kinubi (Wellens 2003: 204)
 úmon áruf gal fĩ difan-á al gi-ǵá
 3PL know COMP EXS guest-PL REL PROG-come
 ‘They know that there are are guests who are coming.’

The use of a *verbum dicendi* as a complementizer resembles the situation in Bari,⁵ where *adi* ‘say’ introduces direct speech (Owens 1997: 163; Miller 2001: 469):

- (22) Bari (Miller 2001: 469)
 mukungu a-kulya adi nan d’ad’ar kakitak merya-mukanat
 sub-chief PAST-say COMP 1SG want worker fifty
 ‘The sub-chief spoke saying: I want fifty workers.’

3.3.4 Lexicon

Since Bari is the main substrate language of Juba Arabic, unsurprisingly it contributes most of its African-derived words: *gúgu* ‘granary’ < Bari *gugu*, *kení* ‘co-wife’ < Bari *köyini*, *lopumég* < Bari *lónyumöng*, *tóngá* ‘pinch’ < Bari *tonga*. In several cases, the Juba Arabic form can be traced to a specific dialect: *d’oñóñ* ‘back of head’ < Pojulu *doñon*, *láña* ‘wander’ < Mundari *lanja* ‘travel’, *nyéte* vs *ñéte* ‘black-eyed pea leaf’ < Bari *nyete* vs Kakwa, Pojulu *ñete*. Moreover, “more Bari lexical items are being borrowed” in Youth Juba Arabic (Nakao 2012: 131): *ka-paparát* ‘butterfly’ < Bari *kapoportat*, *lukulúli* ‘bat’ < Bari *lukululi*. Several other substrate and adstrate languages have contributed to the lexicon of Juba Arabic (Nakao 2012; 2015): *adúngú* ‘harp’ < Acholi *aduñu*; *b’ónǵo* ‘pumpkin’ < Bongo *b’onǵo*; *báfura* ‘cassava’ < Dinka *bafora* ‘manioc, (sweet) cassava’; *káwu* ‘cow-pea’ < Ma’di *kau*; *malangí* < bottle’ < Bangala/Lingala *molangi*; *kámbe* ‘belt’ < Swahili *kamba*; *imbíró* ‘palm tree’ < Zande *mbíró*. Some sixty lexical items found in the earliest records of Ugandan Kinubi can be traced back to various substrate

⁵Unsurprisingly, in Juba Arabic “the use of *adi* as in Bari [is] the most frequent [...] in particular among speakers of Bari origin” (Miller 2001: 470; author’s translation).

8 Arabic pidgins and creoles

languages (Avram 2017a): *lawoti* ‘neighbours’ < Acholi *lawoti* ‘fellow, friend’; *ko-rufu* ‘leaf’ < Bari *korofokorɔpɔ* ‘leaves’; *lwar* ‘abscess’ < Dinka *luär* ‘pain of a swelling’; *seri* ‘fence’ < Lugbara *seri* ‘plant used for fencing’; *mukuta* ‘key’ < Pāri *mukuta*.

The influence of Luganda and Swahili as adstrate languages is already documented in early Ugandan Kinubi (Avram 2017a): Ugandan Kinubi *kibra/kibera* ‘forest’ < Luganda *e-kibira*, *nyinveza* ‘fix’ < Luganda *nyweza* ‘make firm, hold firmly’; *dirisa* ‘window’ < Swahili *dirisha*; *kibanda* ‘shed’ < Swahili *kibanda* ‘small shed’. The lexicon of modern Ugandan Kinubi is characterized by a large number of loanwords from Luganda and Swahili (Wellens 2003; Nakao 2012: 133–134), such as: *mé(é)mvu* ‘banana’ < Luganda *amaemvu* ‘bananas’, *ntulége* ‘zebra’ < Luganda *e-ntulege*; *karibísha* ‘welcome’ < Swahili *karibisha* ‘welcome’, *sangá/šangá* ‘be surprised’ < Swahili *shangaa*. In some cases, these loanwords have replaced previously attested compounds consisting of Arabic-derived elements:⁶ early Ugandan Kinubi *mária bitá murhúm* ‘widow’, lit. ‘wife of the deceased’ vs. modern Ugandan Kinubi *mamwánu* ‘widow’ < Luganda *nnamuwandu*. As for the lexicon of modern Kenyan Kinubi, it is strongly influenced by Swahili. Luffin (2004) lists some 170 loanwords from Swahili (out of approximately 1,400 words recorded), from a wide range of domains, for example: *barabára* ‘highway’ < Swahili *barabara*, *serikáli* ‘government’ < Swahili *serikali*, *tafaúti* ‘difference’ < Swahili *tafauti*, *úza* ‘sell’ < Swahili *ku-uza*. Swahili has also contributed several function words: *badáye* ‘after’ < Swahili *baadaye* ‘afterwards’, *íle* ‘these’ < Swahili *ile*, *na* ‘and, with’ < Swahili *na*. Kenyan Kinubi lexical items have occasionally undergone semantic shift or semantic extension under the influence of the meanings of their Swahili counterparts (Luffin 2014: 315): *destúr* ‘tradition’, cf. Swahili *desturi* ‘tradition’; *fáham* ‘to understand, to remember’, cf. Swahili *-fahamu* ‘to understand, to remember’.

In some cases, the exact origin of loanwords found in Juba Arabic cannot be established: *búra* ‘cat’ < Acholi, Bongo, Dinka, Pāri *bura*, Didinga *buura*; *daná* ‘bow’ < Bari, Jur *dan*, Didinga *d’anga*, Dinka *dhan*; *pondú* ‘cassava leaf’ < Bangala, Kakwa, Lingala *pondu*, Pojulu *pöndu*. The same holds for a number of loanwords attested in early Ugandan Kinubi (Avram 2017a): *bongo* ‘cloth’ < Acholi, Lendu, Lugbara, Zande *bongo*, Bari *bongo*; *godogodo* ‘thin from illness’ < Acholi, Avokaya, Bari, Baka, Lotuho, Moru, Zande *godogodo* ‘very weak, thin, sick(ly)’; *mukungu* ‘headman’ < Acholi, *mukunju*, Bari *mɔkɔŋgɔ*, Lugbara *mukungu*, Luganda *o-mukungu* ‘(sub-) chief’. This is also true of several Kinubi words attested in more recent sources (Wellens 2003; Nakao 2012: 133–134): *júju* ‘shrew’ < Bari

⁶See also Tosco & Manfredi (2013: 509).

Andrei Avram

juju, Ma'di *juju*; *kingilo* 'rhinoceros' < Avokaya *kingili*, Moru *kingile*. In some cases, the occurrence of alternative forms is due to their different SLs: *banğa* 'debt' < Bari *banja*, Lugbara *banja*, Luganda *e-bbanja* vs. *banya* 'debt' < Acholi *banya*.

Under the influence of the substrate and adstrate languages, some Arabic-derived lexical items have undergone semantic extension, thereby becoming polysemous in Juba Arabic (Nakao 2012: 136), e.g. *gówi* 'hard; difficult', cf. Acholi *tek*, Bari *logo*, Lotuho *gol*, Ma'di *okpo*, Swahili *kali*.

Juba Arabic "compensates its lexical gaps through the lexification of Arabic morphosyntactic sequences" (Tosco & Manfredi 2013: 509). A case in point are Juba Arabic compounds, formed via juxtaposition or with their two members linked by the possessive particle *ta* (Manfredi 2014b: 308–309). These include calques after several substrate languages (Nakao 2012: 136), e.g. *ída ta fil* 'elephant trunk', cf. Acholi *ciŋ lyec*, Bari *kōnin lo tome*, Dinka *ciin akɔɔn*, Jur *ciŋ lyec*, Lotuho *naam tome*, Shilluk *bate lyec*, lit. 'arm (of) elephant'. Kinubi also exhibits a number of calques (Nakao 2012; Avram 2017a; Manfredi, this volume). Some of these compounds and phrases can be traced to several SLs, as in the following early Ugandan Kinubi examples (Avram 2017a): *gata kalam* 'decide, judge', cf. Acholi *ŋɔlb kop* 'decide, give judgment', Bongo *ad'oci kudo*, Jur *ŋɔl lubo*, Pări *ŋondi lubo*, lit. 'cut word/speech'; Dinka *wèt tèm* 'decide, give the sentence', lit. 'word cut'; *jua bita ter* 'nest', cf. Acholi *ot winyo*, Bari *kadi-na-kwen*, Belanda Bor *kwot winy*, Shilluk *wot winyo*, Zande *dumô zirê*, lit. 'house (of) bird'. Other calques, presumably more recent ones, reflect the growing influence of Swahili on Kenyan Kinubi (Luffin 2014: 315): *bakán wáy* 'together', cf. Swahili *pamoja* 'together', lit. 'place one', *mára wáy wáy* 'seldom', cf. Swahili *mara moja moja* 'seldom', lit. 'time one one'.

To conclude, SL agentivity accounts for the small number of loanwords and calques recorded in the earliest stage (i.e. pidginization) of Juba Arabic and Kinubi. At a later stage (i.e. after nativization), the larger number of loanwords and calques is a result of borrowing under RL agentivity.

4 Arabic-lexifier pidgins in the Middle East

4.1 Current state and historical development

Several Arabic-lexifiers pidgins have emerged in the Middle East. These include Romanian Pidgin Arabic, Pidgin Madame, Jordanian Pidgin Arabic, and Gulf Pid-

8 *Arabic pidgins and creoles*

gin Arabic. The first three can be classified as work force pidgins.⁷ Gulf Pidgin Arabic also started out as work force pidgin (Smart 1990: 83), but it is now an interethnic contact language (Avram 2014: 13).⁸

Romanian Pidgin Arabic (Avram 2010) was a short-lived pidgin, formerly used on Romanian well sites in Iraq, in locations in the vicinity of Ammara, Basra, Kut, Nassiriya, Rashdiya and Rumaila. Romanian Pidgin Arabic emerged after 1974, when Romanian well sites started operating in Iraq. Romanians typically made up two thirds of the oil crews, with Arabs making up the final third. The first Gulf War and the subsequent withdrawal of the Romanian oil rigs put an end to the use of Romanian Pidgin Arabic.

Immigration of Sri Lankan women to Arabic-speaking countries is reported to have started in 1976 (Bizri 2010: 16), but the large influx into Lebanon came later, in the early 1990s. Pidgin Madame is spoken in Lebanon by Sri Lankan female domestic workers and their Arab employers, mostly in the urban centres of the country.

Jordanian Pidgin Arabic (Al-Salman 2013) is used in the city of Irbid, in the Ar-Ramtha district in the north of Jordan, in interactions between Jordanians and Southeast Asian migrant workers of various linguistic backgrounds. However, only Jordanian Pidgin Arabic as spoken by Bengalis is documented.

Gulf Pidgin Arabic is a blanket term designating the varieties of pidginized Arabic used in Saudi Arabia and the countries on the western coast of the Arab Gulf, i.e. Kuwait, the United Arab Emirates, Oman, Bahrain, and Qatar.

4.2 Contact languages

The main languages involved in the emergence of Romanian Pidgin Arabic are Romanian, Egyptian Arabic (spoken by immigrant workers), and Iraqi Arabic. A small minority of the participants in the language-contact situation had some knowledge of English.

The other pidginized varieties of Arabic in the Middle East share the characteristic of having various Asian languages as their substrate.⁹ For Pidgin Madame, the main contact languages are Lebanese Arabic, as the lexifier language, and Sinhalese. Another language, with a much smaller contribution, is English. In the case of Jordanian Pidgin Arabic, the contact languages are mainly Jordanian

⁷These are pidgins which “came into being in work situations” (Bakker 1995: 28).

⁸That is, one which is “used not just for trade, but also in a wide variety of other domains” (Bakker 1995: 28).

⁹Bizri (2014: 385) therefore suggests the cover term “Asian Migrant Arabic pidgins”.

Andrei Avram

Arabic and Bengali. The contribution of English is very limited. As for Gulf Pidgin Arabic, it emerged in a contact situation of striking complexity. On the one hand, Arabic, the lexifier language, is represented by several dialects, which are not all subsumed under what is known as Gulf Arabic, in spite of what the name of the pidgin suggests. On the other hand, the number of languages spoken by the immigrant workers is staggering: for instance, in the United Arab Emirates the 200 nationalities and 150 ethnic groups speak some 150 languages. Adding to the complexity of the language-contact situation is the fact that these languages are typologically diverse. Last but not least, English also plays a role in interethnic communication, particularly in the service sector.

4.3 Contact-induced changes

4.3.1 Phonology

The phonology of all the pidginized varieties of Arabic in the Middle East exhibits the outcomes of SL agentivity, which also accounts for the occurrence of considerable intra- and inter-speaker variation (Avram 2010: 21–22; Bizri 2014: 393; Avram 2017b: 133).

Consider first Romanian Pidgin Arabic. The following are features characteristic of speakers with Romanian as L1. The pharyngeals are either replaced or deleted: *habib* ‘friend’ < IA/EA *ḥabīb*, *mufta* ‘key’ < IA/EA *muftāḥ*; *saa* ‘hour’ < IA/EA *sāʿa*. Plain consonants are substituted for pharyngealized ones: *halas* ‘ready’ < IA/EA *ḥalāṣ*. Both velar fricatives are replaced: *hamsa* ‘five’ < IA/EA *ḥamsa*; *šogol* ‘work (N)’ < IA *šuy(u)l*. Geminate consonants are degeminated: *sita* ‘six’ < IA/EA *sitta*. There is no distinction between short and long vowels, either in lexical items of Arabic origin or in those from English: *lazim* ‘must’ < IA/EA *lāzim*; *slip* ‘sleep’ < English *sleep*. A feature typical of speakers with Egyptian or Iraqi Arabic as L1 is the substitution of /b/ for Romanian or English /p/ and /v/: *bibul* ‘people, men’ < English *people*; *gib* ‘give, bring’ < English *give*.

Consider next a number of selected features, generally typical of Pidgin Madame, Jordanian Pidgin Arabic, and Gulf Pidgin Arabic. Pharyngeals are either replaced: Pidgin Madame *hareb* ‘war’ < LA *ḥareb*; Jordanian Pidgin Arabic *bisallih* ‘repair’ < JA *biṣalliḥ* ‘repair.IMP.3SG.M’; Gulf Pidgin Arabic *aksan* ‘best’ < GA *aḥsan*, *hut* ‘put’ < GA *ḥutt* ‘put.IMP.2SG.M’; or deleted: Pidgin Madame *ēki* ‘cry’ < LA *əḥki* ‘cry.IMP.2SG.F’; Jordanian Pidgin Arabic *arabi* ‘Arabic’ < JA *ʿarabi*; Gulf Pidgin Arabic *araf* ‘know’ < GA *ʿaraf*. The pharyngealized consonants are replaced by plain counterparts: Pidgin Madame *sarep* ‘envelope’ < LA *ẓaref*; Jordanian Pidgin Arabic *bandora* ‘tomato’ < JA *bandōra*; Gulf Pidgin Arabic *halas* < GA *ḥalāṣ*;

8 Arabic pidgins and creoles

or they are realized as retroflex ones: Pidgin Madame *ṭawīle* ‘long’ < LA *ṭawīle* ‘long.F.SG’. The velar stops are replaced by velar stops or, less frequently, by /h/: Pidgin Madame *sokon* ‘warm’ < LA *suḥun* ‘warm’, *sogol* < LA *šəyəl* ‘work’; Jordanian Pidgin Arabic *kamsa* ‘five’ < JA *ḥamsa*, *sukul* ‘work (N)’ < JA *šuyəl*, *zagīr* ‘small’ < JA *šayīr*; Gulf Pidgin Arabic *kubus* ‘bread’ < GA *ḥubz*, *halas* ‘finish’ < GA *ḥalaš*; *yistokol* ‘work’ < GA *yīštuṭul* ‘work.IMPF.3SG.M’, *šugl* ‘work’ < GA *šuyəl*. Geminate consonants generally undergo degemination (Næss 2008: 36; Avram 2014: 15): Jordanian Pidgin Arabic *sitin* ‘sixty’ < JA *sittīn*; Gulf Pidgin Arabic *sita* ‘six’ < GA *sitta*.

Moreover, consonants not found in the L1s of the users of Gulf Pidgin Arabic may also be replaced. For instance, Indonesian, Javanese, Sinhalese and Tagalog speakers may substitute [p] for /f/: Pidgin Madame *palēpil* ‘falafel’ < LA *falēfil*; Jordanian Pidgin Arabic *pi* ‘in’ < JA *fī* Gulf Pidgin Arabic *napar* ‘person’ < GA *naḥar*; Indonesian and Sinhalese speakers may realize /z/ as [s] or [dʒ]: Pidgin Madame *esa* ‘if’ < LA *iza*; Gulf Pidgin Arabic *sēn* ~ *dʒēn* ‘good’ < GA *zēn* (Bizri 2014: 393; Avram 2017b: 133). Bengali and Sinhalese speakers may replace /š/ with [s]: Pidgin Madame *sū* ‘what’ < LA *šū*; Jordanian Pidgin Arabic *su* ‘what’ < JA *šū*.

Finally, although phonetically long vowels do occur, vowel length is not distinctive, as shown by the occurrence of variation, e.g. Gulf Pidgin Arabic *baden* ~ *badēn* ‘then’ < GA *bašdēn*.

4.3.2 Syntax

There is relatively little that can be attributed to SL agentivity in the syntax of the Arabic-lexifier pidgins in the Middle East (Almoaily 2013; Al-Salman 2013; Avram 2014; Bizri 2014; Avram 2017b; Bakir 2017).

Since the substrate of these varieties, with the exception of Romanian Pidgin Arabic, consists of many SOV languages, e.g. Bengali, Hindi/Urdu, Malayalam, Punjabi, Persian, Sinhalese, Tamil, this word order is occasionally attested (Avram 2017b: 133–134; Bizri 2014: 403). For instance, direct objects may occur in pre-verbal position:

- (23) a. Pidgin Madame (Bizri 2010: 227)
 mīšter kilot sīli
 mister underwear take off
 ‘Mister takes off his underwear.’
 b. Gulf Pidgin Arabic (Avram 2017b: 133)

Andrei Avram

ana čiko sūp
1SG child see
'I will see my children.'

In attributive possession constructions the order of constituents is possessor–possessee:

- (24) a. Pidgin Madame (Bizri 2010: 198)
kullu māmā benet
all mother girl
'All mother's girls.'
- b. Gulf Pidgin Arabic (Næss 2008: 87)
ana jawd bādēn ysīr Jakarta stokol
1SG husband then go Jakarta work
'Then my husband went to work in Jakarta.'

Adjectives generally precedes the nouns they modify:

- (25) Pidgin Madame (Bizri 2010: 119)
bīr bēt
big house
'A big house.'

Similarly, adverbs precede the adjectives they modify:

- (26) a. Pidgin Madame (Bizri 2010: 119)
tīr gūd
very good
'very good'
- b. Gulf Pidgin Arabic (Avram 2014: 25)
sem-sem kalām
same speak
'They speak in the same way.'

Occasional instances of postpositions are attested:

- (27) a. Pidgin Madame (Bizri 2010: 132)
mister mayik masārē
mister with money
'Mister has the money.'

8 Arabic pidgins and creoles

- b. Gulf Pidgin Arabic (Avram 2014: 25)

zamal **fok**

camel above

‘Above the camel.’

Interestingly, Pidgin Madame has a focalized negative copula, derived etymologically from English *no*:

- (28) Pidgin Madame (Bizri 2010: 133)

māmā bīrūt **no**

mother Beirut NEG.FOC

‘It’s not in Beirut that my mother is.’

This resembles the Sinhalese negator *nemiyi*, which “is used only in focalized phrases” (Bizri 2010: 69):

- (29) Pidgin Madame (Bizri 2010: 69)

bat kāve mama **nemeyi**

rice ate 1SG NEG.FOC

‘It is not I who ate the rice.’

4.3.3 Lexicon

Imposition under SL agentivity accounts for the fact that there are few instances of transfer of lexical items from the various SLs into the non-dominant RL (i.e. the pidgin).

The lexicon of Romanian Pidgin Arabic includes words of Romanian and English origin (Avram 2010: 32): *mašina* ‘car’ < Romanian *mașină*, *sonda* ‘oil rig’ < Romanian *sonda*; *spik* ‘speak, say, tell’ < English *speak*, *tumač* ‘much, many’ < English *too much*. Occasionally, non-Arabic words undergo semantic extension under the influence of phonetically similar Arabic words (Avram 2010: 32): *gib* ‘give; bring’ < English *give*, cf. EA *gīb* ‘bring.IMP.2SG.M’.

The lexicon of all the other pidginized varieties of Arabic spoken in the Middle East includes loanwords from English: Pidgin Madame *ambasi* < English *embassy*; *go* < English *go*, *kam* < English *come*, *no gūd* ‘bad’ < English *no good*, *oké* < English *OK*; Jordanian Pidgin Arabic *bēbi* ‘child’ < English *baby*, *finiš* ‘finish’ < English *finish*, *fisa* ‘visa’ < English *visa*; Gulf Pidgin Arabic *hazband* < English *husband*, *pēšent* ‘patient’ < English *patient*. However, as noted by Smart (1990: 113) concerning Gulf Pidgin Arabic, “it is difficult to say [...] whether they are a true part of the pidgin” or rather nonce borrowings.

Andrei Avram

Given the extreme diversity of the substrate, it is not surprising that only a few words from the SLs have made it into the lexicon of Gulf Pidgin Arabic (Avram 2017b: 134–135): *ača* ‘fine’ < Urdu *achā* ‘good, very well’, *ǧaldi/ǧeldi/ǧeldi* < Urdu/Hindi *jaldī* ‘quick’.

Jordanian Pidgin Arabic and Gulf Pidgin Arabic exhibit light verb constructions which may well be calques after Bengali (noun/adjective + *kara* ‘make’) and/or Hindi/Urdu (noun/adjective + *karnā* ‘make’) and/or Persian – noun/adjective + *kardan* ‘make’): Jordanian Pidgin Arabic *sawwi zadīd* ‘renew’, lit. ‘make new’; Gulf Pidgin Arabic *sawwi suāl* ‘ask’, lit. ‘make a question’, *sawwi zalān* ‘upset’, lit. ‘make angry’.

5 Conclusion

This chapter has shown that Arabic-lexifier contact languages emerged primarily through imposition under SL agentivity, in line with the typology of contact languages (Winford 2005: 396; 2008: 128).

The effects of imposition are most obvious in the phonology, syntax and the syntax-semantics interface, and to a lesser extent in the morphology and the lexicon. In the phonology, SL agentivity induces the loss or replacement of certain phonemes not found in the SLs. However, there are also instances of imposition in the sense of transfer from the SLs. As seen, for example, in Turku and Bongor Arabic, some consonants occur only in loanwords from the substrate languages. The occurrence of such loanwords confirms the fact that imposition under SL agentivity may include transfer of lexical items into the RL. Borrowing under RL agentivity has generally played a far less significant role in the development of Arabic pidgins and creoles. As expected, it mostly involves transfer of lexical items; these may lead to the borrowing of certain consonant phonemes, as seen in, for example, Juba Arabic and Kinubi. Finally, borrowing has been shown to include transfer of morphological material as well.

A notable difference between Juba Arabic and Kinubi on the one hand, and the Arabic-lexifiers in the Middle East on the other hand, resides in the relative weight of imposition under SL agentivity and borrowing under RL agentivity. As we have seen, Juba Arabic and Kinubi exhibit the effects of both imposition in their earliest stage (i.e. pidginization), and of borrowing in their latest stage (i.e. nativization). In contrast, imposition is pervasive in the Arabic-lexifier pidgins in the Middle East, given that these varieties have not undergone nativization.

There are still a number of issues awaiting resolution. For instance, the identification of the SLs is rendered difficult by their number and typological diversity.

8 Arabic pidgins and creoles

This difficulty is further compounded by the fact that some substrate languages are still under-researched. This is particularly the case of the substrate languages of Juba Arabic and Kinubi. Also, the distinction between substrate and adstrate languages is blurred (Nakao 2012: 132), particularly when varieties emerge and develop *in situ*, as, for example, with Juba Arabic. Further research also needs to consider the effects of the existence of a creole continuum in Juba Arabic as well as of bilingual and monolingual speakers of the language on the relative importance of restructuring, imposition and borrowing. The extent of restructuring and imposition, for instance, is presumably much greater in basilectal and L2 varieties, as opposed to acrolectal and monolingual varieties of the language. The same holds for Bongor Arabic, which, as shown, appears to be undergoing depidginization. Last but not least, further investigations are necessary to establish whether Gulf Pidgin Arabic is evolving towards stabilization, possibly becoming closer to its lexifier, via borrowing of morphological material or is rather undergoing constant repidginization, essentially via imposition.

Further reading

Miller (1993), Nakao (2012), and Luffin (2014) illustrate in detail substratal and adstratal influence on Juba Arabic and Kinubi.

Avram (2019) analyzes the substratal input in the lexicon of Turku.

Avram (2017b) and Bakir (2017) discuss the various sources of Gulf Pidgin Arabic.

Andrei Avram

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	NEG	neg
ChA	Chadian Arabic	OBJ	object
EA	Egyptian Arabic	PASS	passive
EXS	existential	PAST	past
FOC	focus	PL	plural
GA	Gulf Arabic	POSS	possessive
IA	Iraqi Arabic	PRF	perfect
IMPF	imperfect	PROX	proximal
JA	Jordanian Arabic	REL	relative
LA	Lebanese Arabic	RL	recipient language
lit.	literally	SA	Sudanese Arabic
N	noun	SL	source language
NC	noun class	SG	singular
		V	verb

References

- Al-Salman, Ibrahim. 2013. *Jordanian Pidgin Arabic*. Irbid: Yarmouk University Master's dissertation.
- Almoaily, Mohammad. 2013. *Language variation in gulf Pidgin Arabic*. Newcastle: Newcastle University doctoral dissertation.
- Avram, Andrei A. 2010. An outline of Romanian Pidgin Arabic. *Journal of Language Contact* 3. 19–37.
- Avram, Andrei A. 2014. Immigrant workers and language formation: Gulf Pidgin Arabic. *Lengua y Migración* 6(2). 7–40.
- Avram, Andrei A. 2017a. Evidence of variation in early records of Nubi. Paper presented at the Eleventh Creolistics Workshop “Assessing old assumptions. New insights on the dynamics of languages”, Justus-Giebig University, Giessen, March 22–24.
- Avram, Andrei A. 2017b. Sources of gulf Pidgin Arabic features. In Simone Bettega & Fabio Gasparini (eds.), *Linguistic studies in the Arabian Gulf*, 131–151. Turin: Università di Torino.
- Avram, Andrei A. 2019. Africanisms in turku. In Catherine Miller, Alexandrine Barontini, Marie-Aimée Germanos, Jairo Guerrero & Christophe Pereira (eds.), *Studies on Arabic dialectology and sociolinguistics. Proceedings of the 12th conference of AIDA held in Marseille from May 30th to June 2nd 2017*, 15–24. Aix-en-Provence: Institut de recherches et d'études sur le monde arabe et musulman.

8 Arabic pidgins and creoles

- Bakir, Murtadha J. 2017. GPA genesis: The extent of substratal influence. Paper presented at the Annual Summer Conference of the Society for Pidgin and Creole Linguistics, University of Tampere, 19–22 June.
- Bakker, Peter. 1995. Pidgins. In Jacques Arends, Pieter Muysken & Norval Smith (eds.), *Pidgins and creoles: An introduction*, 130–157. Amsterdam: John Benjamins.
- Bakker, Peter. 2008. Pidgins versus creoles and pidgincreoles. In Silvia Kouwenberg & John Victor Singler (eds.), *The handbook of pidgin and creole studies*, 130–157. Malden, MA: Wiley-Blackwell.
- Bizri, Fida. 2010. *Pidgin Madame: Une grammaire de la servitude*. Paris: Geuthner.
- Bizri, Fida. 2014. Unity and diversity across Asian migrant Arabic pidgins in the middle East. *Journal of Pidgin and Creole Languages* 29(2). 385–409.
- Bureng Vincent, George. 1986. Juba Arabic from a bari perspective. In Gerrit J. Dimmendaal (ed.), *Current approaches to African linguistics*, vol. 3, 71–78. Dordrecht: Foris.
- Cook, Albert R. 1905. *Dinka and Nubi Vocabulary*. Unpublished manuscript.
- De Angelis, Pietro. 2002. *Vocabolario zande*. Trieste: Edizioni Università di Trieste.
- Jenkins, Edward V. 1909. *English–Arabic vocabulary with grammar and phrases representing the language as spoken by the Uganda Sudanese in the Uganda and British East Africa Protectorate*. Kampala: The Uganda Company.
- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Luffin, Xavier. 2004. *Un créole arabe: le kinubi de Mombasa. Etude descriptive*. Brussels: Université Libre de Bruxelles doctoral dissertation.
- Luffin, Xavier. 2005. *Un créole arabe: le kinubi de Mombassa, Kenya*. Munich: Lincom Europa.
- Luffin, Xavier. 2013. Some new information about Bongor Arabic. In Mena Lafkioui (ed.), *African Arabic: Approaches to dialectology*, 161–183. Berlin: De Gruyter Mouton.
- Luffin, Xavier. 2014. The influence of Swahili on kinubi. *Journal of Pidgin and Creole Languages* 29(2). 299–318.
- Manfredi, Stefano. 2014a. Qualche osservazione sull'espressione del plurale nominale nel creole arabo di juba. *RiCognizioni* 2(1). 53–61.
- Manfredi, Stefano. 2014b. Strategie di rilessificazione in juba Arabic. In Alberto Manco & Domenico Silvestri (eds.), *L'etimologia: Atti del XXXV Convegno della Società di Glottologia*, 305–311. Rome: Il Calamo.

Andrei Avram

- Manfredi, Stefano. 2017. *Arabi Juba: Un pidgin-créole du Soudan du Sud*. Leuven: Peeters.
- Manfredi, Stefano. 2018. Arabic as a contact language. In Reem Bassiouney & Elabbas Benmamoun (eds.), *The Routledge handbook of Arabic linguistics*, 407–20. London: Routledge.
- Meldon, J. A. 1913. *English–Arabic dictionary of words and phrases used by the Sudanese in Uganda*. Unpublished manuscript.
- Miller, Catherine. 1989. Bari interference in juba Arabic. In Lionel Bender (ed.), *Proceedings of the 4th Nilo-Saharan Symposium*, 1–10. Hamburg: Helmut Buske.
- Miller, Catherine. 1993. Restructuration morpho-syntaxique en juba-Arabic et ki-Nubi: à propos du débat universaux/substrat et superstrat dans les études créoles. *Matériaux arabes et sudarabiques: Groupe d'études de linguistique et de littératures arabes et sudarabiques* 5. 137–174.
- Miller, Catherine. 2001. Grammaticalisation du verbe 'dire' et subordination en juba Arabic. In Robert Nicolai (ed.), *Leçons d'Afrique: Filiation, rupture et reconstitution des langues. Un hommage à G. Manessy*, 455–482. Leuven: Peeters.
- Moi, Daniel R., Mario L. B. Kuduku, Mary M. Michael, Simon H. John, Raphael Z. Mafoi & Nyoul G. Kuduku. 2014. *Bongo grammar book*. Juba: SIL-South Sudan.
- Muraz, Gaston. 1926. *Vocabulaire du patois arabe tchadien ou "Tourkou" et des dialectes sara-madjinnngaye et sara-m'baye (S.-O. du Tchad)*. Paris: Charles Lavauzelle.
- Næss, Unn Gyda. 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. Oslo: University of Oslo Master's dissertation.
- Nakao, Shuichiro. 2012. Revising the substratal/adstratal influence on Arabic creoles. In Osamu Hieda (ed.), *Challenges in nilotic linguistics: phonology, morphology, and syntax*, 127–49. Tokyo: ILCAA.
- Nakao, Shuichiro. 2015. Another story of bangala: language contact with south sudanese arabic creole. Paper presented at the 8th World Congress of African Linguistics, Kyoto University, 20–24 August.
- Nebel, Arthur. 1979. *Dinka–English, English–Dinka dictionary*. Bologna: Editrice Missionaria Italiana.
- Owen, H. R. & G. J. Keane. 1915. *An abbreviated vocabulary in Hindustani Luganda Lunyoro Swahili Nubi: Designed for the use of the Uganda Medical Service*. Bukalasa: White Father's Printing Press.
- Owens, Jonathan. 1985. The origins of East African Nubi. *Anthropological Linguistics* 27(2). 229–271.

8 Arabic pidgins and creoles

- Owens, Jonathan. 1997. Arabic-based pidgins and creoles. In Sarah G. Thomason (ed.), *Contact languages: A wider perspective*, 125–172. Amsterdam: John Benjamins.
- Pozzati, Aurelio & G. Panza. 1993. *Dizionario Giur.* Trieste: Edizioni Università di Trieste.
- Smart, Jack R. 1990. Pidginization in gulf arabic: a first report. *Anthropological Linguistics* 32(1–2). 83–119.
- Tosco, Mauro & Stefano Manfredi. 2013. Pidgins and creoles. In Jonathan Owens (ed.), *The Oxford handbook of Arabic linguistics*, 496–519. Oxford: Oxford University Press.
- Tosco, Mauro & Jonathan Owens. 1993. Turku: A descriptive comparative study. *Sprache und Geschichte in Afrika* 14. 177–267.
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 1995. Outlining a model of the transmission phenomenon in language contact. *Leuvense Bijdragen* 88. 63–85.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Van Coetsem, Frans. 2003. Topics in contact linguistics. *Leuvense Bijdragen* 92. 27–99.
- Wellens, Inneke. 2003. *An Arabic Creole in Africa*. Nijmegen: Catholic University of Nijmegen doctoral dissertation.
- Winford, Donald. 2005. Contact-induced change: classification and processes. *Diachronica* 22(2). 373–427.
- Winford, Donald. 2008. Processes of creole formation and related contact-induced language change. *Journal of Language Contact* 2(1). 124–145.

Chapter 9

Arabic in the diaspora

Luca D'Anna

Università degli Studi di Napoli 'L'Orientale'

This paper offers an overview of contact-induced change in diasporic Arabic. It provides a socio-historical description of the Arab diaspora, followed by a sociolinguistic profile of Arabic-speaking diasporic communities. Language change is analyzed at the phonological, morphological, syntactic and lexical level, distinguishing between contact-induced change and internal developments caused by reduced input and weakened monitoring. In the course of the description, parallels are drawn between diasporic Arabic and other contemporary or extinct contact varieties, such as Arabic-based pidgins and Andalusí Arabic.

1 Current state and historical development

The terms ARABIC IN THE DIASPORA and ARABIC AS A MINORITY LANGUAGE have been used to designate two distinct linguistic entities, namely Arabic *Sprachinseln* outside the Arabic-speaking world and Arabic in contemporary migration settings. The two situations correspond to the two major social processes that give rise to language contact: conquest and migration. In the former case, speakers of Arabic were isolated from the central area in which the Arabic language is spoken, exposed to a different dominant language and consequently underwent a slow process of language erosion (and eventually shift) usually spanning across several generations. This situation often gives rise to long periods of relatively stable bilingualism, where contact-induced change is more noticeable (Sankoff 2001: 641). In migration contexts, on the contrary, language shift occurs at a faster pace, sometimes within the lifespan of the first generation and usually no later than the third (Canagarajah 2008: 151).

This chapter analyzes contact-induced change in migration contexts. Arab migration to the West started in the late nineteenth century, with the first wave of



Luca D'Anna

migrants who left Greater Syria to settle in the United States and Latin America. The first migrants were mostly Christian unskilled workers, followed by more educated Lebanese, Palestinians, Yemenis and Iraqis after World War II. During the 1950s and 1960s, more migrants continued to settle in the US, while the unstable political situations in Palestine, Lebanon and Iraq resulted in a fourth wave in the 1970s and 1980s (Rouchdy 1992a: 17–18). Because of the events that took place during the last two decades and that resulted in a further destabilization of the entire Middle East, immigration toward the US has never stopped, even though recent American policies have considerably reduced the intake of refugees and immigrants. In 2016, however, 84,995 refugees were resettled in the US, with two Arabic-speaking countries (Syria and Iraq) featuring among the top five states that make 70% of the total intake.¹

Large-scale migration to Western Europe from Arabic-speaking countries began in the wake of the decolonization process during the 1960s and mainly involved speakers from North Africa (Morocco, Algeria and Tunisia). Following a common trend in labor migration, men arrived first, followed by their wives and children. In 1995, a total of 1,110,545 Moroccans, 655,576 Algerians and 279,813 Tunisians lived in Europe, mostly in France, the Netherlands, Belgium, Germany and Italy (Boumans & de Ruiter 2002: 259–260). The socioeconomic profile of the first immigrants mainly consisted of unskilled laborers, usually with low education rates. After six decades from the first wave of immigration, however, most communities consist today of a first, second and third generation, while the political upheaval which started at the end of 2010 resulted in a new wave of young immigrants. Both old and new immigrants had to face the economic crisis that hit Europe in the early 1990s and, again, in 2007, with particularly harsh consequences for the immigrant population (Boumans & de Ruiter 2002: 261).

The sociolinguistic profile of Arabic-speaking communities in the diaspora is quite diverse in different parts of the world and can be analyzed using the ethno-linguistic vitality framework, according to which status, demographics, and institutional support shape the vitality of a linguistic minority (Giles et al. 1977; Ehala 2015). Arabic-speaking immigrants do not usually enjoy a particularly high status, while the level of institutional support is variable. The first waves of immigration to the US, for instance, had to face an environment that was generally hostile to foreign languages. The English-only movement actively worked to impose the exclusive employment of English in public places, while the immigrants themselves committed to learning and using English to integrate into mainstream

¹Data come from the US Department of State. <https://www.state.gov/j/prm/releases/factsheets/2017/266365.htm>, accessed April 2, 2019.

9 *Arabic in the diaspora*

American life. Only in the aftermath of 9/11 did American policymakers begin to re-evaluate the importance of Arabic (and other heritage languages), considering it a resource for homeland security (Albirini 2016: 319–320). Other countries, such as the Netherlands, provided higher levels of formal institutional support, including Arabic in school curricula. These efforts did not achieve the desired goals, however, mostly because the great linguistic diversity of the Moroccan community living in the Netherlands cannot be adequately represented in the teaching curricula. Moroccans in the Netherlands, in fact, speak different Arabic dialects, alongside three main varieties of Berber, namely Tašelhit, Tamazigt and Tarifit (Extra & de Ruiter 1994: 160–161). The voluntary Home Language Instruction, however, provides instruction in Modern Standard Arabic, even though writing skills are only taught starting from 3rd grade (Extra & de Ruiter 1994: 163–165). This is not, of course, the language students are exposed to at home, but attempts to introduce Moroccan dialect or Berber are generally opposed by parents, who value Classical Arabic for its religious and cultural relevance. Similar Home Language Instruction programs are found in most European countries, even though their implementation is sometimes carried out by local governments (in the Netherlands and Germany), private organizations (in Spain) or even by the governments of the origin country (in France) (Boumans & de Ruiter 2002: 264–265). The Italian town of Mazara del Vallo in Sicily represents an extreme case, since the members of the Tunisian community obtained from the Tunisian government the opening of a Tunisian school, where a complete Arabic curriculum is offered and Italian is not even taught as a second language. Until the end of the 1990s, this school, opened in 1981, was the first choice for Tunisian families, who hoped for a possible return to Tunisia. When it eventually became clear that this was unlikely to happen, enrollments consequently declined, which means that Arabic teaching is no longer available to the community in any form (D’Anna 2017b: 73–77). Issues of diglossia and language diversity thus undermine Home Language Instruction programs, which usually occupy a marginal position within school curricula.

Given the generally low status of, and insufficient institutional support for, Arabic-speaking communities in the diaspora, demographic factors are often decisive in determining the ethnolinguistic vitality of the community. While speakers of Arabic are usually scattered in large areas where the dominant language is prevalently spoken, in some Dutch towns Moroccan youth make up 50% of the population of certain neighborhoods (Boumans 2004: 50). At the other end of the continuum, we find closely-knit communities, living in the same neighborhood, such as in Mazara del Vallo, where Tunisians hailing from the two neighbor-

Luca D'Anna

ing towns of Mahdia and Chebba constitute up to 70% of the population of the old town (D'Anna 2017b: 27). All things being equal, given the low status of the Tunisian community and the mediocre institutional support they receive, it is primarily demographic factors which have resulted in the preservation of Arabic in this community beyond the threshold of the third generation.²

In the light of what has been said above, and despite some notable exceptions, Arabic diasporic communities are characterized by relatively rapid processes of language shift, both in the US (Daher 1992: 29) and in Europe (Boumans & de Ruiter 2002: 282). This means that the processes of contact-induced change observed in diasporic communities of Arabic are generally the prelude to language loss. The importance of studying language change in migrant languages, however, also resides in the fact that the same changes usually take place, at a much slower rate, in the standard spoken in the homeland. Internally motivated change in diasporic varieties, from this perspective, often represent an accelerated version of language change in the homeland. Contact-induced change, on the other hand, sometimes suggests parallels with the socially different process of pidginization (Gonzo & Saltarelli 1983: 194–195). The study of Arabic-speaking diasporic communities, thus, can help us shed light on the more general evolution of the language, with regard to both contact-induced and internally-motivated change.

2 Contact languages

Contact languages for diasporic Arabic-speaking communities include, but are not restricted to, American (Rouchdy 1992b) and British English (Abu Haidar 2012), Portuguese in Brazil (Versteegh 2014: 292), French (Boumans & Caubet 2000), Dutch (Boumans 2000; 2004; 2007; Boumans & Caubet 2000; Boumans & de Ruiter 2002), Spanish (Vicente 2005; 2007) and Italian (D'Anna 2017b; 2018). Some contact situations are better described than others, as in the case of English, French and Dutch. At the other end of the continuum, research on the outcome of contact between Italian and Arabic is extremely recent, and data on Portuguese are scarce.

In the following sections, we will draw from the sources so far cited to describe the main phenomena of language change occurring in diasporic Arabic at the phonological, morphological, syntactic and lexical level, highlighting possible parallels with comparable changes in other non-diasporic varieties of Arabic.

²Other factors also played a minor role in the preservation of Arabic in Mazara del Vallo (D'Anna 2017b: 80–81).

3 Contact-induced changes in diasporic Arabic

Despite the great variety of contact languages, it is possible to individuate a number of phenomena that predictably occur in diasporic Arabic-speaking communities. It is not always easy, however, to assess whether an individual phenomenon is due to contact or whether it is, on the contrary, the result of internal development (Romaine 1989: 377). Gonzo & Saltarelli (1977: 177) put the matter as follows:

While it seems clear that some types of changes are due to interference from the dominant language, and others may be attributable to sociological and other external pressures, there are some changes which are language-internal. The latter type is in accordance with a principle of regularization and code reduction which one might expect when the language is acquired in a weakly monitored sociolinguistic environment.

The concept of WEAKENED MONITORING, a situation in which a generally accepted standard and the reinforcement of correct norms are lacking, is an effective tool of analysis when investigating language change in diasporic communities (Gonzo & Saltarelli 1977; 1983). In a situation of weakened monitoring, processes of language change that are occurring slowly in other varieties of the language can be sped up.

In the following sections, interference between languages will be referred to as TRANSFER, which occurs from the SOURCE LANGUAGE (SL) to the RECIPIENT LANGUAGE (RL). If the speaker is dominant in the SL, transfer is more specifically defined as IMPOSITION. If, on the contrary, the speaker is dominant in the RL, transfer is defined as BORROWING (Van Coetsem 1988; 2000; Lucas 2015). While the concept of linguistic DOMINANCE will be extensively used in this paper, one final *caveat* concerns the difficulty of individuating the dominant language (which may actually shift) in second-generation speakers. Lucas identifies a category of 2L1 speakers, who undergo the simultaneous acquisition of two distinct native languages (Lucas 2015: 525). The linguistic trajectory of most second-generation speakers, however, usually involves two consecutive stages in which first the heritage and then the socially dominant language function as the dominant language. While the heritage language is almost exclusively spoken at home during early childhood, in fact, second-generation speakers gradually shift to the socially dominant language when they start school and consequently expand their social network.

Luca D'Anna

3.1 Phonology

In the domain of phonology, diasporic varieties of Arabic generally go in the direction of the loss of marked phonemes (Versteegh 2014: 293). The phonemes undergoing erosion are generally the emphatic and post-velar ones, while the loss is not usually systematic, featuring a great deal of inter and intra-individual variation. In non-diasporic communities, adults, peers and institutions provide corrective feedback to children during their process of language acquisition, while in immigrant communities, due to the weakened monitoring mentioned above, the chain of intergenerational transmission is less secure. Some phenomena of phonetic loss thus have a developmental origin, and are equally common in pidgins and dying languages (Romaine 1989: 372–373). Consider the following example:

- (1) Tunisian Arabic, Mazara del Vallo (D'Anna 2017b: 85)
 ʕala ɥāṭr-i ʕarbi u nnəʒʒəm naʕrəf aktər wāəd
 on thought-OBL.1SG Arab and can.IMP.1SG know.IMP.1SG more one
 mia lingua
 POSS.1SG.F language
 'Because I'm an Arab and I can know above all my language.'

The speaker in sample (1) realizes the voiced pharyngeal fricative /ʕ/, one of the phonemes that are usually lost, but then fails to realize its voiceless counterpart /ħ/ in *wāəd* < *wāħəd* 'one'.³ Similar phenomena also occur, as noted above, in Arabic-based pidgins and creoles, such as Juba Arabic (Manfredi 2017: 17, 21; cf. Avram, this volume).

In the process of phonological erosion, therefore, contact languages seem to have a limited impact. If the dominant language does not feature, in its phonemic inventory, the phoneme that is being eroded, it fails to reinforce whatever input young bilingual speakers receive in the other L1 in the contexts of primary socialization. Reduced input and weakened monitoring, however, play a bigger role, allowing forms usually observed in the earliest stages of language acquisition by monolingual children to survive and spread. It is relatively common, for instance, to observe the presence of shortened or reduced forms, such as *qe* < *lqe* 'he found', *ħal* < *nħal* 'bees', *ləd* < *uləd* 'kid', which sometimes give rise to phenomena of compensation, such as in *uləd* > *ləd* > *lədda* 'kid' (Tunisian

³Similar phenomena of phonetic simplification occur in peripheral varieties of Arabic and *Sprachinseln*, such as Nigerian Arabic (Owens 1993: 19–20; this volume), Cypriot Maronite Arabic (Borg 1985; Walter, this volume), Uzbekistan Arabic (Seeger 2013) and Maltese (Borg & Azzopardi-Alexander 1997: 299; Lucas & Čěplö, this volume). The single varieties here mentioned vary with regard to the phonological simplification they underwent.

diasporic Arabic, Mazara del Vallo, Italy; D'Anna 2017b: 85). In diasporic communities, reduced forms are more easily allowed to survive and spread, occurring in the speech of teenagers, as in the examples reported here. Once again, the same phenomenon also occurs in pidgin and dying languages:

In the case of dying and pidgin languages it may be that children have greater scope to act as norm-makers due to the fact that a great deal of variability exists among the adult community (Romaine 1989: 372–373).

In conclusion, the phonology of diasporic Arabic does not seem to be heavily influenced by borrowing from contact languages. The combined action of reduced input and weakened monitoring, on the other hand, is responsible for the unsystematic loss of marked phonemes and for the survival and spread of reduced forms.

3.2 Morphology

The complex mixture of concatenative and non-concatenative morphology in the domain of Arabic plural formation has been one of the main focuses of research in situations of language contact resulting from migration. Once again, borrowing from contact languages and independent developments occur side by side.

In Arabic, both concatenative and non-concatenative morphology contribute to plural formation. Concatenative morphology, which consists in attaching a suffix to the singular noun, yields the so-called sound plurals, that is, in spoken Arabic, the plural suffixes *-īn* and *-āt* respectively. It has been argued that sound feminine plural is the default plural form according to the morphological underspecification hypothesis, even though masculine is the default gender in all other domains of plural morphology (Albirini & Benmamoun 2014: 855–856). While sound masculine plural is specified for (+human), in fact, sound feminine plural has the semantic feature (\pm human). Non-concatenative, or broken, plurals require a higher cognitive load, since they involve the mapping of a vocalic template onto a consonantal root.⁴ Sound feminine plurals are acquired by children by the age of three, while broken plurals involving geminate and defective roots are not mastered until beyond the age of six (Albirini & Benmamoun 2014: 857–858). After the age of five, however, heritage speakers of Arabic become increasingly exposed to their L2, which encroaches upon their acquisition of broken plurals. It has thus been convincingly demonstrated that heritage speakers display a better command of sound plurals and that, in the domain of broken plurals, some

⁴The notion of root and pattern, which has long been at the core of the morphology of Arabic, has been recently been criticized (Ratchiffe 2013), even though psycholinguistic studies seem to confirm the existence of the root in the mental lexicon of native speakers (Boudelaa 2013).

Luca D'Anna

are more affected by language erosion than others (Albirini & Benmamoun 2014: 858–859). Across different varieties of diasporic Arabic, therefore, plural morphology displays both contact phenomena due to borrowing and internal developments that are akin to what might be called RESTRUCTURING, that is:

changes that a speaker makes to an L2 that are the result not of imposition but of interpreting the L2 input in a way that a child acquiring an L1 would not (Lucas 2015: 525).⁵

Borrowing from the contact languages can take two forms. In rare cases, the suffix plural morpheme of the contact language is directly borrowed, as in the examples *ḥuli-s* ‘sheep-PL’, *ḥmar-s* ‘donkeys’ and *l-ṣud-s* ‘the horses’⁶ collected from one Moroccan informant in the Netherlands (Boumans & de Ruiter 2002: 274). Sometimes, however, transfer works in a subtler way, which consists in the generalization of the sound masculine plural suffix *-īn*,⁷ by analogy with the default form of the contact language, yielding *ḥul-in* ‘sheep-PL’, *ḥmār-in* ‘donkeys’, *ṣewd-in* ‘horses’ (Boumans & de Ruiter 2002: 274). A study conducted by Albirini & Benmamoun (2014: 866–867) shows that L2 learners of Arabic usually tend to overgeneralize the sound masculine plural, wrongly perceived as a default form, while heritage speakers more often resort to the Arabic-specific default, i.e. sound feminine plural. The cases of borrowing reported above, therefore, represent an idiosyncratic exception.

On the other hand, the non-optimal circumstances under which Arabic is learned in diasporic communities often result in overgeneralization processes that cannot be directly attributed to contact. One of them is, as noted above, the generalization of the sound feminine plural *-āt*. In the domain of broken plurals, moreover, not all patterns are equally distributed. The iambic pattern, consisting of a light syllable followed by one with two moras (CVCVVC), is the most common among Arabic broken plurals (Albirini & Benmamoun 2014: 857). As a consequence, it is often generalized by heritage speakers of Levantine varieties (Syrian, Lebanese, Palestinian and Jordanian) living in the US, yielding forms like *fallāḥ* ‘farmer’ pl. *aflāḥ* / *fulūḥ* (target plural *fallāḥ-īn*), *šubbāk* ‘window’ pl. *šubūk* (target plural *šabābīk*), *ṭabbāx* ‘cook’ pl. *ṭabāʔix* (target plural *ṭabbāx-īn*)

⁵In this case, of course, the speaker would not be re-interpreting an L2, but an L1 learned under reduced input conditions and subject to language erosion.

⁶The target form is here *ṣewd-an*, so that also vowel quality is not standard.

⁷The suffix for masculine plural *-īn* is realized with a short vowel in the diasporic Moroccan varieties that are being discussed.

(Albirini & Benmamoun 2014: 865).⁸

Borrowing does not involve plural morphemes only, but other classes as well. In Mazara del Vallo, for instance, young speakers occasionally use the Sicilian diminutive morpheme *-eddru* with Arabic names, creating morphological hybrids of the kind illustrated in (2) :

- (2) Tunisian Arabic, Mazara del Vallo (D'Anna 2017b: 107)
 Grazie safwani-ceddruu⁹
 thanks Safwan-DIM
 'Thanks little Safwan.'

This type of borrowing, quite widespread among young speakers, seems to replicate another instance of contact-induced change that occurred in an extinct variety of Arabic. Andalusí Arabic, in fact, borrowed from Romance the diminutive morpheme *-el* (e.g. *tarabilla* 'mill-clapper' < *tarab+ella* 'little music'), incidentally etymologically cognate with the Sicilian *-eddru* (Latin *-ellum* > Sicilian *-eddru/-eddu*) (*A descriptive and comparative grammar of Andalusí Arabic* 2013: 60). The behavior of the young Tunisian speakers of Mazara del Vallo, who use these Sicilian diminutives in a playful mode, might represent the first stage of the same process that resulted in the transfer of this morpheme into Andalusí Arabic (D'Anna 2017b: 108).

While plurals represent one of the most common areas of change in diasporic Arabic, morpheme borrowing is a much rarer phenomenon, which probably occurs in situations of more pronounced bilingualism. The above two examples, however, provide a representative exemplification of the effect of language contact in the domain of morphology.

3.3 Syntax

Borrowing and restructuring also happen in the domain of syntax. As has been noted both for Moroccans in the Netherlands (de Ruiter 1989: 99) and Tunisians in Italy (personal research), second-generation speakers tend to use simpler clauses

⁸The overgeneralization of some broken plural patterns indicates that the root and pattern system is still productive in heritage speakers, as opposed, for instance, to speakers of Arabic-based pidgins and creoles. Recent studies, however, have advanced the hypothesis that the iambic pattern involves operations below the level of the word, but without necessarily entailing the mapping of a template onto a consonantal root (Albirini et al. 2014: 112).

⁹The utterance appeared as a Facebook post in the timeline of one of my informants and was transcribed verbatim.

Luca D'Anna

than monolingual speakers, namely main or subordinate clauses to which no other clause is attached, as evident from the following sample:

- (3) Tunisian Arabic, Mazara del Vallo (personal research)
 m-baʔd əl-uləyyəd rqað u l-kaləb zāda u l-žrāna
 from-after DEF-boy.DIM sleep.PRF.3SG.M and DEF-dog also and DEF-frog
 ḥaržət mən əl-wāḥəd ēh dabbūsa
 exit.PRF.3SG.F from DEF-one HESIT bottle
 ‘Then the little boy slept and also the dog and the frog escaped from the
 hum bottle.’

Accordingly, they also display the effects of language erosion in establishing long-distance dependencies typical of more complex clauses (Albirini 2016: 305).

Palestinian and Egyptian speakers born in the US have also been found to realize overt pronouns in sentences that opt for the pro-drop strategy in the speech of monolinguals, which is probably due to the influence of English (Albirini et al. 2014: 283). Preliminary observations on second-generation Tunisians in Italy, in fact, do not show the same phenomenon. Since Italian is, like Arabic, a pro-drop language, the use of overt pronouns in American diasporic Arabic can be considered as a case of syntactic borrowing or convergence (Lucas 2015), depending on the speakers’ degree of bilingualism.

The syntax of negation is another area in which language erosion triggers phenomena that seem to be happening, albeit at a slower rate, in non-diasporic communities. Egyptian speakers in the US, for instance, seem to overgeneralize the monopartite negator *miš* / *muš* at the expense of the default discontinuous verbal negator *ma...-š*:

- (4) Egyptian Arabic in the US (Albirini & Benmamoun 2015: 482)
 huwwa miš rāḥ l-kaftiria
 3SG.M NEG go.PRF.3SG.M to-cafeteria
 ‘He didn’t go to the cafeteria.’

Example (4) represents a deviation from the standard Cairene dialect spoken by monolinguals. In Egypt, however, the negative copula *miš* / *muš* represents a pragmatically marked possibility to negate the *b*-imperfect (Brustad 2000: 302), while in Cairo it is now the standard negation for future tense (*miš ḥa-...*, contrasting with *mā-ḥa-...-š* in some areas of Upper Egypt (Brustad 2000: 285). More generally, therefore, *miš* / *muš* is gaining ground at the expense of the discontinuous negation (Brustad 2000: 285), so that what we observe in diasporic Egyptian Arabic might just be an accelerated instance of the same process.

Another major area of language change, documented in most diasporic languages, is the erosion of complex agreement systems (Gonzo & Saltarelli 1983: 192). In diasporic Arabic, heritage speakers show relatively few problems with subject–verb agreement, but struggle with the subtleties of noun–adjective agreement (Albirini et al. 2013: 8). Although verb agreement features more cells to be filled in its paradigm, noun–adjective agreement is less straightforward. Plural nouns, in fact, can trigger adjective agreement in the sound or broken plural or in the feminine singular depending on factors involving humanness, individuation and the morphological shape of both the noun and the adjective, with marked dialectal variation (D’Anna 2017a: 103–104). Heritage speakers thus perform significantly better when default agreement in the masculine singular is required (Albirini et al. 2013: 8), but display evident signs of language erosion when more complex structures are involved:

- (5) Egyptian Arabic in the US (Albirini 2014: 740)
- | | | | | | |
|---------------|-------------------|-------------|------------|---------|----------|
| wi-kamān | baḥibb | arūḥ | l-Detroit | ʔašān | ʕinda-ha |
| and-also | love.IMPF.IND.1SG | go.IMPF.1SG | to-Detroit | because | at-3SG.F |
| maṭāʕim | *mumtaz-īn | | | | |
| restaurant.PL | excellent-PL.M | | | | |
- ‘And I also like to go to Detroit because it has excellent restaurants.’

In (5), the speaker selects the sound masculine plural, while non-human plural nouns require either the broken plural or the feminine singular in Egyptian Arabic. Once again, language change in diasporic Arabic, where the language is learned under reduced input conditions, tends to replicate processes of language change that happened or are happening in the Arabic-speaking world. In the case of agreement, the standardization that the agreement system underwent in the transition from pre-Classical to Classical Arabic has been convincingly explained as emerging from the overgeneralization of frequent patterns by L2 learners (Belnap 1999).

Finally, isolated cases show syntactic borrowing or convergence¹⁰ at the level of word order, which is usually preserved in diasporic contexts:

- (6) Moroccan Arabic in the Netherlands (Boumans 2001: 105)

¹⁰Once again, considering this phenomenon as syntactic borrowing or convergence depends on the speaker’s language dominance, which is not clear from the source and is not easily ascertained in second-generation speakers, whose dominant language is often subject to shift.

Luca D'Anna

u ʔtat l-u dyal-u l-lhem
and give.PRF.3SG.F to-3SG.M GEN-3SG.M DEF-meat
'And she gave it [i.e. the dog] its meat.'

Sample (6) illustrates an extreme case of word order change, in which the possessive *dyal-u* 'its' precedes the head. Overgeneralization of permissible (but sometimes pragmatically marked) word orders, however, occur much more frequently. Egyptian heritage speakers in the US, for instance, use SVO order 77.65% of the time, vs 52.64% for Egyptian native speakers (Albirini et al. 2011: 280–281).

In situations of stable bilingualism, such as in some Arabic *Sprachinseln*, convergence with contact languages can result in permanent alterations to word order. In Buxari Arabic, for instance, transitive verbs feature a mandatory SOV word order, with optional resumptive pronoun after the verb. Cleft sentences such as the following one are quite common in all Arabic dialects:

- (7) Egyptian Arabic (Ratcliffe 2005: 145)
il-fustān gībt-u
DEF-dress get.PRF.1SG-3SG.M
'I got the dress.'

In Bukhari Arabic, which has long been in contact with SOV languages (such as Persian and Tajik), this structure became the standard for transitive verbs, so that the resumptive pronoun can also be dropped, as in the following sample:

- (8) Bukhari Arabic (Ratcliffe 2005: 144)
fāt ʔūd ḥada
INDEF stick take.PRF.3SG.M
'He took a stick.'

3.4 Lexicon

In the domain of lexical borrowing, which has attracted considerable interest among scholars, the situation of bilingualism in diasporic contexts poses some methodological issues in the individuation of actual loanwords. The production of heritage speakers, in fact, is inevitably marked by frequent phenomena of codeswitching, which makes difficult to distinguish between nonce-borrowings (Poplack 1980) and codeswitching. If we define lexical borrowing as “the diachronic process by which languages enhance their vocabulary” (Matras 2009: 106), in fact, it is not clear which language is here enhancing its vocabulary, since diasporic

varieties of Arabic are not discrete varieties and feature the highest degree of internal variability. A possible solution to this impasse consists in looking exclusively at the linguistic properties of the alleged loanword. In this vein, Adalar & Tagliamonte (1998: 156) have shown that, when foreign-origin nouns appear in contexts in which they are completely surrounded by the other language, they are treated like borrowings (in this case, nonce-borrowings) at the phonological, morphological and syntactic level. When, on the contrary, they appear in bilingual (or multilingual) utterances, they represent cases of codeswitching, patterning with the language of their etymology. The domain of lexical borrowing in diasporic varieties of Arabic, however, is an area that needs further research.

4 Conclusion

This chapter has offered an overview of the main phenomena of contact-induced change observed in Arabic diasporic communities, distinguishing them from internal developments due to reduced input and weakened monitoring. Diasporic communities rarely feature situations of stable bilingualism, so that language change usually corresponds to language attrition and is followed by the complete shift to the dominant language. The study of language change in diasporic communities, however, constitutes an interesting field of investigation, both in itself and for the insight it can give us into language change in monolingual communities. Change at the phonological, morphological and syntactic level finds parallels in comparable phenomena that have occurred in the history of Arabic (such as in the case of agreement) or that are occurring as we speak (such as in the case of the spread of the negator *miš* in Egyptian Arabic). Not by chance, similar phenomena also occur(red) in the Arabic-based pidgins of East Africa, such as Juba Arabic. Various scholars, in fact, have maintained that the mechanisms of change differ in the degree of intensity, but not in their intrinsic nature, from those operating in less extreme situations of contact (e.g. Miller 2003: 8; Lucas 2015: 528).

On the other hand, the analysis of contact phenomena in diasporic communities poses some methodological issues with regard to the categories of borrowing, imposition and convergence (Van Coetsem 1988; 2000). These categories, in fact, imply the possibility to define clearly the speaker's dominant language or, at least, to define him as a stable 2L1 speaker. This is rarely the case with heritage speakers, whose repertoires follow trajectories in which language dominance shifts, usually from the heritage language to the socially dominant one. This process is usually concomitant with the beginning of school education, but we lack

Luca D’Anna

theoretical and methodological tools to determine with accuracy the speaker’s position on the trajectory.

Further avenues of research on this topic thus include a more rigorous investigation of emerging and shifting repertoires and the analysis of the complex relation between diasporic languages, pidginization and creolization, which has already been the object of a number of contributions (e.g. [Gonzo & Saltarelli 1983](#); [Romaine 1989](#)).

Further reading

[Rouchdy \(1992b\)](#) is the first description of Arabic in the US.

[Rouchdy \(2002\)](#) analyzes more broadly language contact and conflict, with a section devoted to Arabic in the diaspora.

[Owens \(2000\)](#) collects essays on Arabic as a minority language, focusing on both *Spracheninseln* and diasporic Arabic, but introducing also historical and cross-ethnic perspectives.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	INDEF	indefinite
DEF	definite article	M	masculine
DIM	diminutive	NEG	negative
F	feminine	OBL	oblique
HESIT	hesitation	PL	plural
GEN	genitive	POSS	possessive pronoun
IMPF	imperfect	PRF	perfect
IND	indicative	SG	singular

Acknowledgments

I am grateful to the University of Mississippi, which generously funded this research and my fieldwork in Mazara del Vallo. To Adam Benkato, for reading the manuscript and providing, as always, his valuable feedback. To all my informants in Mazara del Vallo, whose patience during the interviews was only matched by their warm hospitality.

References

- A descriptive and comparative grammar of Andalusí Arabic*. 2013. Leiden: Brill.
- Abu Haidar, Farida. 2012. Arabic and English in conflict: Iraqis in the UK. In Aleya Rouchdy (ed.), *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, 286–297. London: Routledge.
- Adalar, Nevin & Sali Tagliamonte. 1998. Borrowed nouns, bilingual people: the case of the “Londrali” in Northern Cyprus. *International Journal of Bilingualism* 2. 139–159.
- Albirini, Abdulkafi. 2014. Toward understanding the variability in the language proficiencies of Arabic heritage speakers. *International Journal of Bilingualism* 18(6). 730–765.
- Albirini, Abdulkafi. 2016. *Modern Arabic sociolinguistics: Diglossia, variation, codeswitching, attitudes and identity*. London: Routledge.
- Albirini, Abdulkafi & Elabbas Benmamoun. 2014. Concatenative and nonconcatenative plural formation in l1, l2, and heritage speakers of Arabic. *The Modern Language Journal* 98(3). 854–872.
- Albirini, Abdulkafi & Elabbas Benmamoun. 2015. Factors affecting the retention of sentential negation in heritage Egyptian Arabic. *Bilingualism: Language and Cognition* 18(3). 470–489.
- Albirini, Abdulkafi, Elabbas Benmamoun & Brahim Chakrani. 2013. Gender and number agreement in the oral production of Arabic heritage speakers. *Bilingualism: Language and Cognition* 16(1). 1–18.
- Albirini, Abdulkafi, Elabbas Benmamoun, Silvina Montrul & Eman Saadah. 2014. Arabic plurals and root and pattern morphology in Palestinian and Egyptian heritage speakers. *Linguistic Approaches to Bilingualism* 4(1). 89–123.
- Albirini, Abdulkafi, Elabbas Benmamoun & Eman Saadah. 2011. Grammatical features of Egyptian and Palestinian Arabic heritage speakers’ oral production. *Studies in Second Language Acquisition* 33(2). 273–303.
- Belnap, R. Kirk. 1999. A new perspective on the history of arabic variation in marking agreement with plural heads. *Folia Linguistica* 33(1–2). 169–186.
- Borg, Albert & Marie Azzopardi-Alexander. 1997. *Maltese*. London: Routledge.
- Borg, Alexander. 1985. *Cypriot Arabic: a historical and comparative investigation into the phonology and morphology of the arabic vernacular spoken by the Maronites of Kormakiti village in the Kyrenia district of north-west Cyprus*. Stuttgart: Deutsche Morgenländische Gesellschaft.
- Boudelaa, Sami. 2013. Psycholinguistics. In *The Oxford handbook of Arabic Linguistics*, 369–392. Oxford: Oxford University Press.

Luca D'Anna

- Boumans, Louis. 2000. Topic pronouns in monolingual Arabic and in codeswitching. In Manwel Mifsud (ed.), *Proceedings of the Third International Conference of l'Association Internationale de Dialectologie Arabe held in Malta 29 March – 2 April 1998*, 89–95. Malta.
- Boumans, Louis. 2001. Moroccan Arabic and Dutch: languages of Moroccan youth in the Netherlands. *Langues et Linguistique: Revue Internationale de Linguistique* 8. 97–120.
- Boumans, Louis. 2004. L'arabe marocaine de la generation ayant grandi aux Pays-Bas. In Dominique Caubet, Jacqueline Billiez, Thierry Bulot, Isabelle Léglise & Catherine Miller (eds.), *Parlers jeunes ici et là-bas: Pratiques et representations*, 49–68. Paris: L'Harmattan.
- Boumans, Louis. 2007. The periphrastic bilingual verb construction as a marker of intense language contact: Evidence from Greek, Portuguese and Maghribian Arabic. In Everhard Ditters & Motzki Harald (eds.), *Approaches to Arabic linguistics presented to Kees Versteegh on the occasion of his sixtieth birthday*, 291–311. Leiden: Brill.
- Boumans, Louis & Dominique Caubet. 2000. Modelling intrasentential codeswitching: a comparative study of Algerian/French in Algeria and Moroccan/Dutch in the Netherlands. In Jonathan Owens (ed.), *Arabic as a minority language*, 113–181. Berlin: Mouton de Gruyter.
- Boumans, Louis & Jan Jaap de Ruiter. 2002. Moroccan Arabic in the European diaspora. In Aleya Rouchdy (ed.), *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, 259–285. London: Routledge.
- Brustad, Kristen. 2000. *The syntax of spoken Arabic: a comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Georgetown: Georgetown University Press.
- Canagarajah, Suresh A. 2008. Language shift and the family: Questions from the Sri Lankan Tamil diaspora. *Journal of Sociolinguistics* 12. 143–176.
- D'Anna, Luca. 2017a. Agreement with plural controllers in Fezzānī Arabic. *Folia Orientalia* 54. 101–123.
- D'Anna, Luca. 2017b. *Italiano, siciliano e arabo in contatto: profilo sociolinguistico della comunità tunisina di Mazara del Vallo*. Palermo: Centro di studi filologici e linguistici siciliani.
- D'Anna, Luca. 2018. Dialect contact in the Tunisian diaspora: Chebba speakers in Mazara del Vallo (Sicily). In *Perspectives on arabic linguistics XXXI: papers from the annual symposium on arabic linguistics, norman, oklahoma, 2017*. Amsterdam: John Benjamins.

- Daher, Nazih Y. 1992. A Lebanese dialect in Cleveland: Language attrition in progress. In *The Arabic language in America*, 25–36. Detroit: Wayne State University Press.
- de Ruiter, Jan Jaap. 1989. *Young Moroccans in the Netherlands: an integral approach to their language situation and acquisition of Dutch*. Utrecht: Rijksuniversiteit doctoral dissertation.
- Ehala, Martin. 2015. Ethnolinguistic vitality. In *The international encyclopedia of language and social interaction*. Hoboken, NJ: Wiley-Blackwell.
- Extra, Guus & Jan Jaap de Ruiter. 1994. The sociolinguistic status of the Moroccan community in the Netherlands. *Indian Journal of Applied Linguistics* 20(1–2). 151–176.
- Giles, Howard, Richard Y. Bourhis & Donald M. Taylor. 1977. Towards a theory of language in ethnic group relations. In Howard Giles (ed.), *Language, ethnicity and intergroup relations*, 307–348. London: Academic Press.
- Gonzo, Susan & Mario Saltairelli. 1977. Migrant languages: Linguistic change in progress. In Marcel De Greve & Eddy Rossel (eds.), *Problèmes linguistiques des enfants des travailleurs migrants*, 167–187. Brussels: Didier AIMAV.
- Gonzo, Susan & Mario Saltairelli. 1983. Pidginization and linguistic change in emigrant languages. In Roger Andersen (ed.), *Pidginization and creolization as language acquisition*, 181–198. Rowley: Newbury House.
- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Manfredi, Stefano. 2017. *Arabi ĵuba: Un pidgin-créole du Soudan du Sud*. Leuven: Peeters.
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Miller, Catherine. 2003. The relevance of Arabic-based pidgins-creoles for Arabic linguistics. *Al-Lugha* 4. 7–46.
- Owens, Jonathan. 1993. *A grammar of Nigerian Arabic*. Wiesbaden: Harrassowitz.
- Owens, Jonathan (ed.). 2000. *Arabic as a minority language*. Berlin: Mouton de Gruyter.
- Poplack, Shana. 1980. Sometimes i'll start a sentence in spanish y termino en español. *Linguistics* 18. 581–618.
- Ratcliffe, Robert. 2005. Bukhara arabic: a metatypized dialect of arabic in central asia. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: case studies from iranian, semitic and turkic*, 141–51. London: Routledge.

Luca D'Anna

- Ratcliffe, Robert R. 2013. Morphology. In Jonathan Owens (ed.), *The Oxford handbook of Arabic linguistics*, 71–92. Oxford: Oxford University Press.
- Romaine, Suzanne. 1989. Pidgins, creoles, immigrant and dying languages. In Nancy C. Dorian (ed.), *Investigating obsolescence: Studies in language contraction and death*, 369–385. Cambridge: Cambridge University Press.
- Rouchdy, Aleya. 1992a. Introduction. In Aleya Rouchdy (ed.), *The Arabic language in America*, 17–21. Detroit: Wayne State University Press.
- Rouchdy, Aleya (ed.). 1992b. *The Arabic language in America*. Detroit: Wayne State University Press.
- Rouchdy, Aleya (ed.). 2002. *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*. London: Routledge.
- Sankoff, Gillian. 2001. Linguistic outcomes of language contact. In Jack Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 638–68. Oxford: Blackwell.
- Seeger, Ulrich. 2013. Zum verhältnis der zentralasiatischen arabischen dialekte. In Renaud Kutty, Ulrich Seeger & Shabo Talay (eds.), *Nicht nur mit engelszungen: beiträge zur semitischen dialektologie. festschrift für werner arnold zum 60. geburtstag*, 313–322. Wiesbaden: Harrassowitz.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Versteegh, Kees. 2014. *The Arabic language*. Second edition. Edinburgh: Edinburgh University Press.
- Vicente, Ángeles. 2005. *Ceuta: une ville entre deux langues. (une étude sociolinguistique de sa communauté musulmane)*. Paris: L'Harmattan.
- Vicente, Ángeles. 2007. Two cases of Moroccan Arabic in the diaspora. In Catherine Miller, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.), *Arabic in the city: Studies in dialect contact and language variation*, 123–145. London: Routledge.

Part II

Language change through contact with Arabic

Chapter 10

Beja

Martine Vanhove

LLACAN (CNRS, INALCO)

This chapter argues for two types of outcomes of the long-standing and intense contact situation between Beja and Arabic in Sudan: borrowings at the phonological, syntactic and lexical levels, and convergence at the morphological level.

1 Current state and historical development

1.1 Historical development of Beja

Beja is the sole language of the Northern Cushitic branch of the Afroasiatic phylum. Recent archeological discoveries show growing evidence that Beja is related to the extinct languages of the Medjay (from which the ethnonym Beja is derived; [Rilly 2014](#): 1175), and Blemmye tribes, first attested on Egyptian inscriptions of the Twelfth Dynasty for the former, and on a Napatan stela of the late seventh century BCE for the latter. For recent discussions, see [Browne \(2003\)](#); [El-Sayed \(2011\)](#); [Zibeliuss-Chen \(2014\)](#); [Rilly \(2014\)](#); and [Rilly \(2018\)](#). The Medjays were nomads living in the eastern Nubian Desert, between the first and second cataracts of the River Nile. The Blemmyes invaded and took part in defeating the Meroitic kingdom, fought against the Romans up to the Sinai, and ruled Nubia from Talmis (modern Kalabsha, between Luxor and Aswan) for a few decades, before being defeated themselves by the Noubades around 450 CE ([Rilly 2018](#)). In late antiquity, the linguistic situation involved, in northern Lower Nubia, Cushitic languages, Northern Eastern Sudanic languages, to which Meroitic and Nubian belong, also Coptic and Greek to some extent, and in the south, Ethio-Semitic. It is likely that there was mutual influence to an extent that is difficult to disentangle today.



Martine Vanhove

1.2 Current situation of Beja

The Beja territory has shrunk a lot since late antiquity, and Beja (*biḍawije:t*) is mainly spoken today in the Red Sea and Kassala States in eastern Sudan, in the dry lands between the Red Sea and the Atbara River. The 1993 census, the last one to include a language question, recorded some 1,100,000 Beja speakers, and there is probably at least double that figure today. There are also some 60,000 speakers in northern Eritrea, and there may be still a few speakers left in Egypt, in the Nile valley at Aswan and Daraw, and along the coast towards Marsa Alam (Morin 1995; Wedekind 2012). In Sudan today, Beja speakers have also settled in Khartoum and cities in central and western Sudan (Hamid Ahmed 2005a: 67).

All Bejas today are Muslims. They consider themselves bedouins, and call themselves *arab* ‘Arab’;¹ they call the ethnical Arabs *balawje:t*. Before the introduction of modern means of transportation, they were traditionally the holders of the caravan trade in the desert towards the west, south and north of their territory, and they still move between summer and winter pastures with their cattle. They also produce sorghum and millet for daily consumption, and fruits and vegetables in the oases. The arrival of Rashaida migrants from Saudi Arabia in the nineteenth century created tensions in an area with meagre resources, but the first contemporary important social changes took place during the British mandate with the agricultural development of the Gash and Tokar areas, and the settlement of non-Beja farmers. The droughts of the mid-1980s brought about a massive exodus towards the cities, notably Port Sudan and Kassala, followed by job diversification, and increased access to education in Arabic, although not generalized, especially for girls, who rarely go beyond primary level (Hamid Ahmed 2005a).

Beja is mostly an oral language. In Eritrea, a Latin script was introduced in schools after independence in 1993, but in Sudan no education in Beja exists. Attempts made by the Summer Institute of Linguistics and at the University of the Red Sea to implement an Arabic-based script did not come to fruition. On the other hand, in the last few years school teachers in rural areas have begun to talk more and more in Beja in order to fight illiteracy (in Arabic) and absenteeism (Onour 2015).

¹In Sudan the term *ʕarab* is widely used for referring to nomad groups in general, and not only to ethnically defined Arabs. Thanks to Stefano Manfredi for this information..

2 Arabic–Beja contact

Contact between Bejas and Arabs started as early as the beginning of islamization, and through trade relations with Muslim Egypt, as well as Arab incursions in search of gold and emerald. Evidence of these contacts lie in the early arabicization of Beja anthroponyms (Záhořík 2007). The date of the beginning of Islamization differs according to authors, but it seems it started as early as the tenth century, and slowly expanded until it became the sole religion between the eighteenth and nineteenth centuries (Záhořík 2007).

We have no information concerning the beginning and spread of Beja–Arabic bilingualism. It is thus often impossible to figure out if a transfer occurred through Beja-dominant speakers or was imposed by fluent Beja–Arabic bilingual speakers, and consequently to tell apart whether a contact-induced feature belongs to the borrowing or to the imposition type of transfer as advocated by Van Coetsem (1988; 2000) and his followers. What is certain though, is that socio-historical as well as linguistic evidence speaks in favour of Beja–Arabic bilingualism as an ancient phenomenon, but in unknown proportion among the population. With the spread of Islam since the Middle Ages, contact with Arabic became more and more prevalent in Sudan.

In this country, which will be the focus of this chapter, bilingualism with Sudanese Arabic is frequent, particularly for men, and expanding, including among women in cities and villages, but to a lesser extent. Bejas in Port Sudan are also in contact with varieties of Yemeni Arabic. Rural Bejas recently settled at the periphery of the big cities have the reputation of being more monolingual than others, which was still the case fifteen years ago (Vanhove 2003).

The Beja language is an integral part of the social and cultural identity of the people, but it is not a necessary component. Tribes and clans that have switched to Arabic, or Tigre, such as the Beni Amer, are considered Bejas. Beja is prestigious, since it allows its speakers to uphold the ethical values of the society, and is considered to be aesthetically pleasing due to its allusive character. The attitude towards Arabic is ambivalent. It is perceived as taboo-less, and thus contrary to the rules of honour, nevertheless it is possible to use it without transgressing them. Arabic is also prestigious because it is the language of social promotion and modernity (Hamid Ahmed 2005b). Language attitudes are rapidly changing, and there is some concern among the Beja diaspora about the future of the Beja language, even though it cannot be considered to be endangered. Some parents avoid speaking Beja to their children, for fear that it would interfere with their learning of Arabic at school, leaving to the grandparents the transmission of Beja

Martine Vanhove

(Wedekind 2012; Vanhove 2017). But there is no reliable quantitative or qualitative sociolinguistic study of this phenomenon. Codeswitching between Beja and Arabic is spreading but understudied.

This sketchy description of the sociolinguistic situation of Beja speaks for at least two types of transfer: (i) borrowing, where the agents of transfer are dominant in the recipient language (Beja); (ii) convergence phenomena, since the difference in linguistic dominance between the languages of the bilingual speakers tends to be really small (at least among male speakers today, and probably earlier in the history of Beja; see Van Coetsem 1988: 87). Imposition has probably also occurred of course, but it is not always easy to prove.

3 Contact-induced changes in Beja

3.1 Phonology

The few contact-induced changes in Beja phonology belong to the borrowing type.

The phonological system of Beja counts 21 consonantal phonemes, presented in Table 1.

Table 1: Beja consonants

	<i>Bilabial</i>	<i>Labio-dental</i>	<i>Alveolar</i>	<i>Retroflex</i>	<i>Palatal</i>	<i>Velar</i>	<i>Labio-velar</i>	<i>Laryngeal</i>
Plosive		f	t	ɖ		k	k ^w	ʔ
	b		d	ɖ		g	g ^w	
Affricate					ɟ			
Fricative			s		ʃ			h
Nasal	m		n					
Trill			r					
Lateral			l					
Approximant	w				j			

The voiced post-alveolar affricate ɟ (often realised as a voiced palatal plosive [ɟ] as in Sudanese Arabic) deserves attention as a possible outcome of contact

with Arabic. Since Reinisch (1893: 17), it is usually believed that this affricate is only present in Arabic loanwords and is not a phoneme (Roper 1928; Hudson 1976; Morin 1995). The existence of a number of minimal pairs in word-initial position invalidates the latter analysis: *ɟi:k* ‘rooster’ ~ *ʃi:k* ‘chewing tobacco’; *ɟhar* ‘chance’ ~ *dhar* ‘bless’; *ɟaw* ‘quarrel’ ~ *ɗaw* ‘jungle’ ~ *ʃaw* ‘pregnancy’ ~ *gaw* ‘house’ (Vanhove 2017). As for the former claim, there are actually a few lexical items such as *bʔaɟi* ‘bed’, *gʷʔaɟi* ‘one-eyed’ (*gʷʔad* ‘two eyes’), that cannot be traced back to Arabic (the latter is pan-Cushitic; Blažek ms. [2000]). Nevertheless, it is the case that most items containing this phoneme do come from (or through) (Sudanese) Arabic: *a:laɟ* ‘tease’, *aɟi:n* ‘dough’, *aɟib* ‘please’, *ʔaɟala* ‘bicycle’, *ʔiɟir* ‘divine reward’, *ɟa:hil* ‘small child’, *ɟabana* ‘coffee’, *ɟalla:j* ‘because of’, *ɟallab* ‘fish’, *ɟanna* ‘paradise’, *ɟanta:ji* ‘djinn’, *ɟarika:n* ‘jerrycan’, *ɟe:b* ‘pocket’, *ɟhali* ‘coal’, *ɟimʔa* ‘week’, *ɟins* ‘sort’, *ɟuwwa* ‘inside’, *ʃaɟil* ‘morning’, *ʃinɟa:n* ‘cup’, *hanɟar* ‘dagger’, *hiɟ* ‘pilgrimage’, *maɟaʔa* ‘famine’, *maɟlis* ‘reconciliation meeting’, *siɟin* ‘prison’, *tarɟima:l* ‘translator’, *waɟɟa* ‘appointment’, and *xawaɟa* ‘foreigner’. It is clear that *ɟ* is not marginal anymore. However *ɟ* is unstable: it has several dialectal variants, *ʃ*, *g* and *d*, and may alternate with the dental *d* or retroflex *ɖ*, in the original Beja lexicon (*ɟiʷʔo:r*/*ɖiʷʔo:r* ‘honourable man’) as well as in loanwords (*aɟi:n*/*aɖi:n* ‘dough’) (Vanhove & Hamid Ahmed 2011; Vanhove 2017). In my data, which counts some 50 male and female speakers of all age groups, this is rarely the case, meaning that there is a good chance that this originally marginal phoneme will live on under the influence of (Sudanese) Arabic.

There are two other consonants in Arabic loanwords that are regularly used by the Beja speakers: *z* and *x*, neither of which can be considered phonemes since there are no minimal pairs.

Blažek (2007: 130) established a regular correspondence between Beja *d* and Proto-East-Cushitic **z*. In contemporary Beja *z* only occurs in recent loanwords from Sudanese Arabic such as *ɟaza* ‘wage’, *ɟo:z* ‘pair’, *rizg* ‘job’, *wazʔ* ‘offer’, *xazna* ‘treasure’, *zama:n* ‘time’, *zirʔa* ‘field’, *zu:r* ‘visit’. It may alternate with *d*, even within the speech of the same speaker as free variants, e.g. *dama:n*, *dirʔa*, *du:r*. The fricative alveolar pronunciation is more frequent among city dwellers, who are more often bilingual. It is difficult to ascertain whether Beja is in the process of re-acquiring the voiced fricative through contact with Sudanese Arabic, or whether it will undergo the same evolution to a dental stop as in the past.

A few recent Arabic loanwords may also retain the voiceless velar fricative *x* (see also Manfredi et al. 2015: 304–305): *xazna* ‘treasure’, *xawaɟa* ‘foreigner’, *xadda:m* ‘servant’, *xa:tar* ‘be dangerous’, *a:xar* ‘last’. In my data, this is usually the

Martine Vanhove

case in the speech of fluent bilingual speakers. We thus have here a probable imposition type of transfer. In older borrowings, even among these speakers, Arabic *x* shifted to *h* (*xajma* > *he:ma* ‘tent’). It may be because these older loans spread in a community which was at that time composed mainly of Beja-dominant speakers, but we have no means of proving this hypothesis.

3.2 Morphology

3.2.1 General remarks

Most Cushitic languages only have concatenative morphology, the stem and pattern schema being at best highly marginal (Cohen 1988: 256). In addition to Beja, Afar and Saho (Lowland East-Cushitic branch), Beja’s geographically closest sisters, are exceptions, and all three languages use also non-concatenative morphology. In Afar and Saho it is far less pervasive than in Beja; in particular they do not use vocalic alternation for verbal derivation, this feature being restricted to the verb flexion of a minority of underived verbs.

Even though Beja and Arabic share a similar type of morphology, the following overview shows that each language has developed its own system. Although they have been in contact for centuries, neither small-scale nor massive borrowing from Arabic morphological patterns can be postulated for the Beja data. An interpretation in terms of a convergence phenomenon is more relevant, both in terms of semantics and forms.

Non-concatenative morphology concerns an important portion of the lexicon: a large part of the verb morphology (conjugations, verb derivations, verbal noun derivations), and part of the noun morphology (adjectives, nouns, “internal” plurals, and to a lesser extent, place and instrument names). In what follows, I build on Vanhove (2012), and Vanhove (2017), correcting some inaccuracies.

3.2.2 Verb morphology

Only one of the two Beja verb classes, the one conjugated with prefixes (or infixes), belongs to non-concatenative morphology. This verb class (V1) is formed of a stem which undergoes ablaut varying with tense–aspect–mood (TAM), person and number, to which prefixed personal indices for all TAMs are added (plural and gender morphemes are also suffixes). V1 is diachronically the oldest pattern, which survives only in a few other Cushitic languages. In Beja V1s are the majority (57%), as against approximately 30% in Afar and Saho, and only five verbs in Somali and South Agaw (Cohen 1988: 256). Table 2 provides examples in the

perfective and imperfective for bi-consonantal and tri-consonantal roots.

Table 2: Perfective and imperfective patterns

	Bi-consonantal <i>dif</i> ‘go’	Tri-consonantal <i>kitim</i> ‘arrive’
PFV	<i>i-dif</i> ‘he went’	<i>i-ktim</i> ‘he arrived’
	<i>i-dif-na</i> ‘they went’	<i>i-ktim-na</i> ‘they arrived’
IPFV	<i>i-n-di:f</i> ‘he goes’	<i>k<an>ti:m</i> ‘he arrives’
	<i>e:-dif-na</i> ‘they go’	<i>e:-katim-na</i> ‘they arrive’

Prefix conjugations are used in Arabic varieties and South Semitic languages but their functions and origins are different. In South Semitic, the prefix conjugation has an aspectual value of imperfective, while in Cushitic it marks a particular morphological verb class. The Cushitic prefix conjugation (in the singular) goes back to auxiliary verbs meaning ‘say’ or ‘be’, while the prefix conjugation of South Semitic has various origins, none of them including a verb ‘say’ or ‘be’ (Cohen 1984). Although different grammaticalization chains took place in the two branches of Afroasiatic, this suggests that the root-and-pattern system might have already been robust in Beja at an ancient stage of the language. It is noteworthy that there are at least traces of vocalic alternation between the perfective and the imperfective in all Cushitic branches (Cohen 1984: 88–102), thus reinforcing the hypothesis of an ancient root-and-pattern schema in Beja. In what proportion this schema was entrenched in the morphology of the proto-Cushitic lexicon is impossible to decide.

Verb derivation of V1s is also largely non-concatenative. Beja is the only Cushitic language which uses qualitative ablaut in the stem for the formation of semantic and voice derivation. The ablaut can combine with prefixes.

Table 3 presents the five verb derivation patterns with ablaut, and Table 4 shows the absence of correspondence between the Beja and Arabic (Classical and Sudanese) patterns. Sudanese patterns are extracted from Bergman (2002: 32–34), who does not provide semantic values.

Among the Semitic languages, an intensive pattern similar to the Beja one is only known in some Modern South Arabian languages spoken in eastern Yemen (not in contact with Beja), where it is also used for causation and transitivity (Simeone-Senelle 2011: 1091). The Modern South Arabian languages are close relatives of Ethio-Semitic languages and it is usually considered that the latter were brought to the Horn of Africa by South-Arabian speakers (Ullendorff 1955). However, this ablaut pattern was not retained in Ethio-Semitic. It is also unknown in

Martine Vanhove

Table 3: V1 derivation patterns with ablaut

	Monosyllabic V1	Plurisyllabic V1
INT	<i>bo:s</i> (<i>bis</i> ‘burry’)	<i>ka:tim</i> (<i>kitim</i> ‘arrive’)
MID	<i>faf</i> (<i>fif</i> ‘pour’)	<i>rimad</i> (<i>rimid</i> ‘avenge’)
PASS	<i>a:to:-ma:n</i> (<i>min</i> ‘shave’)	<i>at-daba:l</i> (<i>dibil</i> ‘gather’) <i>am-he:jid</i> (<i>ha:jid</i> ‘sew’)
RECP	<i>amo:-ga:d</i> (<i>gid</i> ‘throw’)	<i>am-gara:m</i> (<i>girim</i> ‘be inimical’)
CAUS		<i>si-katim</i> (<i>kitim</i> ‘arrive’)

Table 4: Comparison between Beja and Arabic derivation patterns

	Beja Plurisyllabic V1	Classical Arabic	Sudanese Arabic
INT	Ca:Ca <i>ka:tim</i> < <i>kitim</i> ‘arrive’	CaCCaCa, Ca:CaCa	CaCCaC, Ca:CaC
MID	CiCaC <i>rimad</i> < <i>rimid</i> ‘avenge’	CuCiCa, iC<t>aCaCa, ta-CaCCaCa, ta-Ca:CaCa, in-CaCaCa, tas-CaCaCa, ista-CaCaCa	it-CaCCaC, it-CaCaC
PASS	at-CaCa:C <i>at-daba:l</i> < <i>dibil</i> ‘gather’ am-Ce:CaC <i>am-he:jid</i> < <i>ha:jid</i> ‘sew’	CuCiCa, iC<t>aCaCa, ta-CaCCaCa, ta-Ca:CaCa, in-CaCaCa, ista-CaCaCa	it-Ca:CaC, it-CaCaC, CiCiC, in-CaCaC
RECP	am-CaCa:C <i>am-gara:m</i> < <i>girim</i> ‘be inimical’	Ca:CaCa , ta-Ca:CaCa, iC<t>aCaCa	Ca:CaC, it-Ca:CaC
CAUS	si-CaCiC <i>si-katim</i> < <i>kitim</i> ‘arrive’	CaCCaCa, ?a-CCaCa	CaCCaC, a-CCaC

Cushitic. In Classical Arabic, the plurisyllabic pattern does not have an intensive value, but a goal or sometimes reciprocal meaning.

Beja is the sole Cushitic language which differentiates between active and middle voices by means of vocalic alternation. Remnants of this pattern exist in some Semitic languages, among them Arabic, in a fossilized form.

In Cushitic, qualitative ablaut for the passive voice only occurs in Beja. Passive formation through ablaut exists in Classical and Sudanese Arabic, but with different vowels. Bergman (2002: 34) mentions that “a handful of verbs in S[udanese] A[rabic]” can be formed this way. For Stefano Manfredi (personal communication) it is a productive pattern in this Arabic variety.

Like the passive voice, the reciprocal is characterized by a qualitative ablaut in *a*: in the stem, but the prefix is different and consists of *am(o:-)*. *m* is not used for verbal derivation in Arabic, which uses the same ablaut, but for the first vowel of disyllabic stems, to express, marginally, the reciprocal of the base form. Most often the reciprocal meaning is expressed by other forms with the *t*-prefixed or infix to the derived form or the base form. In some other Cushitic languages *-m* is used as a suffix for passive or middle voice (without ablaut). In Beja *m*- can also marginally be used as a passive marker, together with ablaut, for a few transitive intensive verbs: *ame:-saj* ‘be flayed’, *ame:-biḏan* ‘be forgotten’.

Although a suffix *-s* (not a prefix as in Beja) is common in Cushitic, Beja is once more the only Cushitic language which uses ablaut for the causative derived form. Neither ablaut nor the *s*-prefix exist in Arabic. Arabic uses different patterns for the causative: the same as the intensive one, i.e. with a geminated second root consonant, and the (ʔ)a-CcaC(a) pattern.

This brief overview shows that Beja has not borrowed patterns from (Sudanese) Arabic, but has at best similar, but not exact, cognate patterns which are marginal in both Classical and Sudanese Arabic.

Beja also has four non-finite verb forms. The simultaneity converb of V1s is the only one with non-concatenative morphology. The affirmative converb is marked for both verb classes with a suffix *-e*: added to the stem: *gid* ‘throw’, *gid-e*: ‘while throwing’; *kitim* ‘arrive’, *kitim-e*: ‘while arriving’. In the negative, the negative particle *ba:-* precedes the stem, and V1s undergo ablaut in the stem (Ci:C and CaCi:C), and drop the suffix; it has a privative meaning: *ba:-gi:d* ‘without throwing’; *ba:-kati:m* ‘without arriving’. No similar patterns exist in Arabic or other Cushitic languages.

Martine Vanhove

3.2.3 Verbal noun derivation

In the verbal domain, non-concatenative morphology concerns only V1s. With nouns, it applies to action nouns (*maṣḍars*) and agent nouns.

There are several *maṣḍar* patterns, with or without a prefix, with or without ablaut, depending mostly on the syllabic structure of the verb. The most frequent ones with ablaut are presented below.

The pattern *m(i(:)/a)-CV(:)C* applies to the majority of monosyllabic verbs. The stem vowel varies and is not predictable: *di* ‘say’, *mi-ja:d* ‘saying’; *dir* ‘kill’, *ma-dar* ‘killing’; *sʔa* ‘sit down’, *ma-sʔa:* ‘sitting’; *ak* ‘become’, *mi-kti* ‘becoming’; *hiw* ‘give’, *mi-jaw* ‘gift, act of giving’. A few disyllabic V1s comply to this pattern: *rik^wij* ‘fear’, *mi-rk^wa:j* ‘fearing’; *jiwid* ‘curl’, *mi:-wad* ‘curling’. Some V1s of the CiC pattern have a Ca:C pattern for *masdars*, without a prefix: *gid* ‘throw’, *ga:d* ‘throwing’. In Classical Arabic, the marginal *maṣḍars* with a prefix concern tri-syllabic verbs, none showing a long vowel in the stem or the prefix, nor a vowel *i* in the prefix.

CiCiC and HaCiC² disyllabic verbs form their *masdars* by vocalic ablaut to *u:*: *kitim* ‘arrive’, *kitu:m* ‘arriving’; *ʔabik* ‘take’, *ʔabu:k* ‘taking’; *hamir* ‘be poor’, *hamur* ‘being poor’. CiCaC V1s, and those ending in *-j*, undergo vocalic ablaut to *e:*: *dig^wag^w* ‘catch up’, *dig^we:g^w* ‘catching up’; *biḍa:j* ‘yawn’, *biḍe:j* ‘yawning’. In Classical Arabic, the *maṣḍar* pattern with *u:* has a different vowel in the first syllable, *a* (in Beja *a* is conditioned by the initial laryngeal consonant), and it is limited almost exclusively to verbs expressing movements and body positions (Blachère & Gaudefroy-Demombynes 1975: 81).

Bergman (2002: 35) provides no information about verbal nouns of the base form in Sudanese Arabic except that they are “not predictable”.

As for agent nouns of V1s, they most often combine ablaut with the suffix *-a:na*, the same suffix as the one used to form agent nouns of V2 verbs, whose stems do not undergo ablaut. The ablaut pattern is the same as with the verbal intensive derivation: *bir* ‘snatch’, *bo:r-a:na* ‘snatcher’; *gid* ‘throw’, *ge:d-a:na* ‘thrower, a good shot’; *dibil* ‘pick up’, *da:bl-ana* ‘one who picks up’. Some tri-consonantal stems have a suffix *-i* instead of *-a:na*: *fibib* ‘look at’, *fa:bb-i* ‘guard, sentinel. Some have both suffixes: *kitim* ‘arrive’, *ka:tm-a:na/ka:tim-i* ‘newcomer’.

These patterns are unknown in Arabic.

²Where H stand for the laryngeals ʔ and *h*.

3.2.4 Noun morphology

3.2.4.1 General remarks

The existence of verbal noun derivation patterns and nominal plural patterns are well recognized in the literature about Beja morphology; for a recent overview, see [Appleyard \(2007\)](#). It is far from being the case for adjective and noun patterns. All noun and adjective patterns linked to V1s are listed below. [Vanhove \(2012\)](#) provides an overview of these patterns which are summed up below.

3.2.4.2 Adjective patterns

There are eight adjective patterns, two of which are shared with nouns. Most are derived from V1 verbs, but the reverse is also attested. A corresponding verb form is inexistent in a few cases. All patterns are based on ablaut, in two cases with an additional suffix *-a*, or gemination of the medial consonant. Arabic has no dedicated adjective pattern (but the active participle pattern of the verbal base form Ca:CaC may express properties). Table 5 provides the full list of patterns with examples. It is remarkable that none of them is similar to those of Classical Arabic or colloquial Sudanese Arabic ([Bergman 2002](#): 17).

Table 5: Adjective patterns

Pattern	Adjective	Verb form
aCa:C	<i>ama:g</i> ‘bad’	<i>mig</i> ‘do evil’
CaCCa	<i>mar?a</i> ‘wide’	<i>mir?</i> ‘be wide’
Ca:Ca(C)	<i>na:k^w</i> ‘short’	<i>nik^w</i> ‘be short’
	<i>da:ji</i> ‘good’	Ø
CaCa(C)	<i>dawil</i> ‘close’	<i>diwil</i> ‘be close’
CaCa:C3	<i>tak^wa:k^w</i> ‘prepared’	<i>tik^wik^w</i> ‘prepare’
CaCa:C-a	<i>raga:g-a</i> ‘long’	<i>rigig</i> ‘stand up’
CiCa:C	<i>fik^wa:n</i> ‘aromatic’	<i>fik^wan</i> ‘emit pleasant odour’
CaCCiC	<i>fallik</i> ‘few’	<i>filik</i> ‘be few’

3.2.4.3 Nouns

There are eleven basic noun patterns related to V1 verbs. Most of the patterns for triconsonantal roots resemble those of Arabic (but are not strictly identical), a

Martine Vanhove

coincidence which is not surprising since both languages have a limited number of vowels. Table 6 provides the full list of these patterns. The CaCi and CiCi patterns are shared with adjectives. The CiCi(C) pattern does not undergo ablaut.

Table 6: Noun patterns

Pattern	Noun	Verb form
CaC	<i>nak^w</i> ‘pregnancy’	<i>nik^wi</i> ‘become pregnant’
CiCa	<i>nisa</i> ‘advise’	<i>nisa</i> ‘advice’
CaCi	<i>sari</i> ‘wakefulness’	<i>sir</i> ‘keep awake’
CaCa	<i>nada</i> ‘dew’	<i>nidaj</i> ‘sweat, exude water’
CiCi(C)	<i>mirʔi</i> ‘width’	<i>mirʔ</i> ‘be wide’
	<i>rifid</i> ‘wealth’	<i>rifid</i> ‘raise, care for cultivation or cattle’
CaCi:C	<i>ʃaɖi:ɖ</i> ‘strip’	<i>ʃiɖiɖ</i> ‘strip off’
	<i>ʃak^wi:n</i> ‘fragrance’	<i>ʃik^wan</i> ‘emit pleasant odour’
CaCi:C-a	<i>rafi:d-a</i> ‘cattle’	<i>rifid</i> ‘raise, care for cultivation or cattle’
Ci:Ca:C	<i>ti:la:l</i> ‘stride’	<i>tilil</i> ‘stride far away from home’
CaCo:C	<i>tabo:k</i> ‘double handful’	<i>tibo:k</i> ‘fill scoop with two hands joined’
CiCu:C-a	<i>tilu:l-a</i> ‘exile’	<i>tilil</i> ‘stride far away from home’
CaCo:C	<i>tabo:k</i> ‘double-handful’	<i>tibo:k</i> ‘fill scoop with cupped hands’
CiCu:C-a	<i>tilu:l-a</i> ‘exile’	<i>tilil</i> ‘stride far away from home’

3.2.4.4 Nouns with prefix *m(V)*-

A few other semantic types of nouns, mostly instrument and place names, are formed through ablaut and a prefix *m(V)*-, like in Arabic. Contrary to Arabic where these patterns are productive, they are frozen forms in Beja (some are not loanwords from Arabic, see the last three examples): *ʔafi* ‘prevent, secure’, *m-ʔafaj* ‘nail, rivet, fastener’; *himi* ‘cover’, *m-himme:j* ‘blanket’; *ginif* ‘kneel’, *mi-gnaf* ‘camp’; *mo:k* ‘take shelter’, *ma-k^wa* ‘shelter’; *rifif* ‘drag an object on ground’, *mi-rfaf* ‘reptile’.

3.2.5 Plural patterns

The so-called “internal plural” patterns are common and frequent in Arabic (and Ethio-Semitic). Beja also has a limited set of internal plural patterns, but it has developed its own system. Ablaut patterns for plural formation mainly concern non-derived nouns containing either a long vowel or ending in a diphthong. Both *i:* and *u:* turn to *i* in the plural, and *a:*, *e:* and *o:* turn to *a*, sometimes with the addition of the plural suffix *-a*; nouns ending in *-aj* turn to a long vowel *-e:j*: *ang^wi:l*, pl. *ang^wil* ‘ear’; *lu:l*, pl. *lil* ‘rope’; *asu:l*, pl. *asil* ‘blister’; *hasa:l*, pl. *hasal/hasal-a* ‘bridle’; *me:k*, pl. *mak* ‘donkey’; *bo:k*, pl. *bak* ‘he-goat’; *ganaj*, pl. *gane:j* ‘gazelle’ (Vanhove 2017).

Even though internal plurals can be considered as a genetic feature, the fact that they are very rare or absent in other Cushitic languages (Zaborski 1986) speaks for a possible influence of Arabic (in Sudan) upon Beja.

3.3 Syntax

3.3.1 General remarks

As far as we know, there are no syntactic calques from Arabic in Beja. There are nevertheless a few lexical and grammatical borrowed items that gave rise to constructions concerning coordination and subordination.

3.3.2 Coordination

One of the three devices that mark coordination is borrowed from Arabic *wa*. It is only used for noun phrases or nominalized clauses (deranked, temporal and relative clauses), whereas the Arabic source particle can be used with noun phrases and simple sentences. *wa* is preposed to the coordinated element in Arabic, but in Beja it is an enclitic particle =*wa*, a position in line with the favoured SOV word order. =*wa* follows each of the coordinated elements. (1) illustrates the coordination of two noun phrases.

- (1) Beja (BEJ_MV_NARR_01_shelter_057)³
 bʔaɖaɖ=wa i=ko:lej=wa sallam-ja=aj=he:b
 sword=COORD DEF.M=stick=COORD give-PFV.3SG.M=CSL=OBJ.1SG
 ‘Since he had given me a sword and the stick...’

³The sources of the examples are accessible online at <http://corporan.huma-num.fr/Archives/corpus.php>; the indications in parenthesis refer to the texts they are extracted from.

Adversative coordination between two simple clauses is also expressed with a borrowing from Arabic: *la:kin* ‘but’.

3.3.3 Subordination

The reason conjunction *sabbi* ‘because’ is a borrowing from the Arabic noun *sabab* ‘reason’. Like most balanced adverbial clauses, it is based on one of the relative clause types, the one nominalized with the noun *na* ‘thing’ in the genitive case. *sabbi* functions as the head of the relative clause.

- (6) Beja (03_camel_192)
 ?akir-a qab qa:b-i:n=e:=na:-ji sabbi
 be_strong-CVB.MNR run.AC run-AOR.3PL=REL=thing-GEN because
 ‘Because it was running so fast..’

sabbi can also be used after a noun or a pronoun in the genitive case: *ombarijo:k sabbi* ‘because of you’.

Terminative adverbial clauses are expressed with a borrowing from Arabic, *hadi:d* ‘limit’. Again the borrowing is the head of the relative clause.

- (7) Beja (BEJ_MV_NARR_51_camel_stallion_026-030)
 o:n i=ka:m=o:k he:=he:b
 PROX.SG.M.ACC DEF.M=camel=POSS.2SG.ACC give[.IMP.SG.M]=OBJ.1SG
 i-ndi e:n baru:k o:=bu:n
 3SG.M-say.IPFV DM 2SG.M.NOM DEF.SG.M.ACC=coffee
 g^w?a-ti=e:b hadi:d
 drink-AOR.2SG.M=REL.M until
 ‘Leave your camel with me, he says, until you have drunk your coffee!’

hadi:d can also be used as a postposition after a noun, in which case it can be abbreviated to *had:* *faɟil-had* ‘until morning’.

3.4 Lexicon

The study of the Beja lexicon lacks research on the adaptation of Arabic loan-words and their chronological layers. There is no statistic on the proportions of lexical items borrowed from Arabic or Ethio-Semitic as compared to those inherited from Cushitic, not to mention Afroasiatic as a whole or borrowed from Nilo-Saharan. Phonetic and morphological changes are bound to have blurred the etymological data, but what is certain is that massive lexical borrowings from

Martine Vanhove

Arabic for all word categories took place at different periods of time, and that the process is still going on. Lexicostatistical studies (Cohen 1988: 267; Blažek 1997) have shown that Beja shares only 20% of basic vocabulary with its closest relatives, Afar, Saho and Agaw.

In this section I mainly concentrate on verbs, because they are often believed to be less easily borrowed in language contact situations (see Wohlgemuth 2009 for an overview of the literature on this topic), which obviously is not the case for Beja.

Cohen (1988) mentions that tri-consonantal V1s contain a majority of Semitic borrowings. I conducted a search of Reinisch's (1895) dictionary, the only one to mention possible correspondences with Semitic languages. It provided a total of 225 V1s, out of which only nine have no Semitic cognates (four are cognates with Cushitic, one is borrowed from Nubian, and one cognate with Egyptian). Even if some of Reinisch's comparisons are dubious, the overall picture is still in favour of massive borrowings from Semitic (96%). It is not easy to disentangle whether the source is an Ethio-Semitic language or Arabic, but until a more detailed study can be undertaken, the following can be said: 55 verbs (20%) have cognates only in Ethio-Semitic (Tigre, Tigrinya, Amharic, and/or Ge'ez); out of the remaining 161 (72%), 85 are attested only in Arabic, 76 also in Ethio-Semitic. Because of the long-standing contact with Arabic for a large majority of Beja speakers in Sudan, and the marginality of contact with Tigre limited to the south of the Beja domain, it is tempting to assume that almost 3/4 of the 76 verbs are of Arabic origin. They may have been borrowed at an unknown time when the new suffix conjugation was still marginal. However, there are also tri-consonantal verbs (V2) which are conjugated with suffixes, albeit less numerous: 164. 141 have cognates with Semitic languages (95 Arabic, 31 Ethio-Semitic, and 15 attested in both branches), six are pan-Cushitic, one is pan-Afroasiatic, one Nubian, six are of dubious origin, and nine occur only in Beja. Does this mean that these borrowings occurred later than for V1s? In the current state of our knowledge of the historical development of Beja, it is not possible to answer this question.

On the other hand, Cohen (1988: 256), in his count of consonants per stem in eight Cushitic and Omotic languages, showed that biconsonantal stems are predominant in six of the languages. By contrast, they form 52.8% of the 770 Beja stems in Roper's (1928) lexicon, and 42.7% of the 611 Agaw stems, almost on a par with bi-consonantal stems (42.2%). What this shows is that massive borrowings from Arabic (or from Ethio-Semitic for Agaw) helped to preserve tri-consonantal stems, which still form a majority of the stems in Beja, unlike in other Cushitic languages.

4 Conclusion

This overview has shown that massive lexical borrowings from Arabic in Beja have helped to significantly entrench non-concatenative morphology in this language. Whether this is a preservation of an old Cushitic system, or a more important development of this structure than in other Cushitic languages under the influence of Arabic, is open to debate, but what is certain is that it is not incidental that this system is so pervasive in Beja, the only Cushitic language to have had a long history of intense language contact with Arabic, the Semitic language where non-concatenative morphology is the most developed. What is important to recall is that Beja non-concatenative morphology shows no borrowings of Arabic patterns (unlike in Modern South Arabian languages; see Bettega and Gasparini, this volume), leading to the conclusion that we are dealing with a convergence phenomenon. Lexical borrowings and morphological convergence are not paralleled in the phonological and syntactic domains where Arabic influence seems marginal.

Much remains to be done concerning language contact between Beja and Arabic, and we lack reliable sociolinguistic studies in this domain. We also lack a comprehensive historical investigation of the Beja lexicon, as well as a sufficiently elaborated theory of phonetic correspondences for Cushitic (1988: 267). Even though important progress has been made, in particular for Beja in the comparison of its consonant system with other Cushitic languages and concerning the etymology of lexical items in some semantic fields, thanks to Blažek (ms. [2000]; 2003a; 2003b; 2006a; 2006b), the absence of a theory of lexical borrowings in Beja (and other Cushitic languages) is still an impediment for a major breakthrough in the understanding of language contact between Beja and Arabic.

Further reading

Apart from Vanhove (2012) on non-concatenative morphology already summarized in §3.2, we lack studies on contact-induced changes in Beja. Vanhove (2003) is a brief article on code-switching in one tale and two jokes based on conversational analysis. Wedekind (2012) is an appraisal of the changing sociolinguistic situation of Beja in Egypt, Sudan and Eritrea.

Martine Vanhove

Acknowledgements

My thanks are due to my Sudanese consultants and collaborators, in particular Ahmed Abdallah Mohamed-Tahir and his family in Sinkat, Mohamed-Tahir Hamid Ahmed in Khartoum, Yacine Ahmed Hamid and his family who hosted me in Khartoum. I am grateful to the two editors of this volume, and wish to acknowledge the financial support of LLACAN, the ANR projects CorpAfroAs and CorTypo (Principal Investigator Amina Mettouchi). This work was also partially supported by a public grant overseen by the French National Research Agency (ANR) as part of the program “Investissements d’Avenir” (reference: ANR-10-LABX-0083). It contributes to the IdEx Université de Paris - ANR-18-IDEX-0001.

Abbreviations

AC	action noun, <i>maṣḍar</i>	LOC	locative
ACC	accusative	M	masculine
AOR	aorist	MID	middle
BCE	before Common Era	MNR	manner
CAUS	causative	NEG	negation
CE	Common Era	NOM	nominative
COM	comitative	OBJ	object
COORD	coordination	OPT	optative
CSL	causal	PASS	passive
CVB	converb	PFV	perfective
DEF	definite	PL	plural
DM	discourse marker	POSS	possessive
F	feminine	PROX	proximal
GEN	genitive	RECP	reciprocal
GNRL	general	REL	relator
IMP	imperative	SG	singular
INDF	indefinite	SMLT	simultaneity
INT	intensive	TAM	tense–aspect–mood
IPFV	imperfective		

References

Appleyard, David. 2007. Beja morphology. In Alan S. Kaye (ed.), *Morphologies of Asia and Africa*, 447–480. Winona Lake, IN: Eisenbrauns.

- Bergman, Elizabeth. 2002. *Spoken Sudanese Arabic: Grammar, dialogues, and glossary*. Springfield, VA: Dunwoody Press.
- Blachère, Régis & Maurice Gaudetroy-Demombynes. 1975. *Grammaire de l'arabe classique*. Third edition. Paris: Maisonneuve et Larose.
- Blažek, Václav. 1997. Cushitic lexicostatistics: The second attempt. In *Afroasiatica italiana*, 171–188. Napoli: Instituto Universitario Orientale.
- Blažek, Václav. 2003a. Beja kinship and social terminology. In Monika R. M. Hasitzka-Johannes & Diethart-Günther Dembski (eds.), *Das alte Ägypten und seine Nachbarn: Festschrift zum 65. Geburtstag von Helmut Satzinger*, 307–340. Krems: Österreichisches Literaturforum.
- Blažek, Václav. 2003b. Fauna in beja lexicon: a fragment of a comparative-etymological dictionary of beja. In Leonid Kogan & Alexander Militarev (eds.), *Studia Semitica: Festschrift for Alexander Militarev*, 230–294. Moscow: Russian State University for Humanities.
- Blažek, Václav. 2006a. Natural phenomena, time and geographical terminology in beja lexicon: fragment of a comparative and etymological dictionary of beja: i. *Babel und Bibel (Memoriae Igor M. Diakonoff)* 2. 365–407.
- Blažek, Václav. 2006b. Natural phenomena, time and geographical terminology in beja lexicon: fragment of a comparative and etymological dictionary of beja: II. *Babel und Bibel* 3. 383–428.
- Blažek, Václav. 2007. Beja historical phonology: Consonantism. In Azeb Amha, Graziano Savà & Maarten Mous (eds.), *Omotic and Cushitic language studies: Papers from the fourth Cushitic Omotic conference, Leiden, 10–12 april 2003*, 125–156. Cologne: Rüdiger Köppe.
- Blažek, Václav. ms. [2000]. Fragment of a comparative and etymological dictionary of beja: anatomical lexicon. Unpublished manuscript.
- Browne, Gerald M. 2003. *Textus Blemmycus: Aetatis Christianae*. Champaign, IL: Stipes.
- Cohen, David. 1984. *La phrase nominale et l'évolution du système verbal en sémitique: Etude de syntaxe historique*. Leuven: Peeters.
- Cohen, David. 1988. Bédja. In David Cohen (ed.), *Les langues dans le monde ancien et moderne*, vol. 3: *Langues chamito-sémitiques*, 270–277. Paris: Editions du CNRS.
- El-Sayed, Rafed. 2011. *Afrikanischstämmiger Lehnwortschatz im älteren Ägyptisch: Untersuchungen zur ägyptisch-afrikanischen lexikalischen Interferenz im dritten und zweiten Jahrtausend v. Chr.* Leuven: Peeters.
- Hamid Ahmed, Mohamed-Tahir. 2005a. “Paroles d’hommes nonorables”: *Essai d’anthropologie poétique des Bedja du Soudan*. Leuven: Peeters.

Martine Vanhove

- Hamid Ahmed, Mohamed-Tahir. 2005b. Ethics and oral poetry in beja society (Sudan). In Catherine Miller (ed.), *Land, ethnicity and political legitimacy in eastern Sudan: Kassala and Gedaref States*, 475–504. Lawrenceville, NJ: Africa World Press/The Red Sea Press.
- Hudson, Richard A. 1976. Beja. In Lionel Bender (ed.), *The non-Semitic languages of Ethiopia*, 97–132. East Lansing, MI: African Studies Center, Southern Illinois University.
- Manfredi, Stefano, Marie-Claude Simeone-Senelle & Mauro Tosco. 2015. Language contact, borrowing and codeswitching. In Amina Mettouchi, Martine Vanhove & Dominique Caubet (eds.), *Corpus-based studies of lesser-described languages: the CorpAfroAs corpus of spoken AfroAsiatic languages*, 283–308. Amsterdam: John Benjamins.
- Morin, Didier. 1995. “Des paroles douces comme la soie”: Introduction aux contes dans l’aire couchitique (bedja, afar, saho, somali). Leuven: Peeters.
- Onour, ʿAbdallah. 2015. *ʿAsbāb al-ʿummiyya fīnd al-Biḡā: dirāsāt ḥāla, wilāyat Kasala, maḥalliyyat Šamāl ad-Daltā, manṭiqat Wagar* [The reasons for illiteracy among the Bejas: Case study, Kassala State, Shamal al-Dalta region, Wagar locality]. Khartoum: University of Khartoum doctoral dissertation.
- Reinisch, Leo. 1893. *Die Beḍauye-Sprache in Nordost-Afrika*. Vol. 1–2. Vienna: Buchhändler der kaiserlichen Akademie der Wissenschaften.
- Reinisch, Leo. 1895. *Wörterbuch der Beḍawiye-Sprache*. Vienna: Alfred Hölder.
- Rilly, Claude. 2014. Language and ethnicity in ancient Sudan. In Julie R. Anderson & Derek A. Welsby (eds.), *The Fourth Cataract and beyond: Proceedings of the 12th International Conference for Nubian Studies*, 1169–1188. Leuven: Peeters.
- Rilly, Claude. 2018. Languages of ancient nubia. In Dietrich Raue (ed.), *Handbook of Ancient Nubia*, 129–152. Berlin: De Gruyter.
- Roper, E. M. 1928. *Tu Beḍawie: An elementary handbook for the use of Sudan government officials*. Hertford: Stephen Austin.
- Simeone-Senelle, Marie-Claude. 2011. Modern South arabian. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1073–1113. Berlin: De Gruyter Mouton.
- Ullendorf, Edward. 1955. *The Semitic languages of Ethiopia: A comparative phonology*. London: Taylor’s.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.

- Vanhove, Martine. 2003. Bilinguisme et alternance bedja-arabe au soudan. In Ignacio Ferrando & Juan José Sánchez Sandoval (eds.), *AIDA 5th conference proceedings, cádiz, september 2002*, 131–142. Cádiz: Universidad de Cádiz, Servicio de Publicaciones.
- Vanhove, Martine. 2012. Roots and patterns in Beja (Cushitic): The issue of language contact with Arabic. In Martine Vanhove, Thomas Stolz, Hitmoi Otsuka & Aina Urdze (eds.), *Morphologies in contact*, 311–326. Berlin: Akademie Verlag.
- Vanhove, Martine. 2016. The manner converb in Beja (Cushitic) and its reinitiation. In Claudine Chamoreau & Zarina Estrada-Fernández (eds.), *Finiteness and nominalization*, 323–344. Amsterdam: John Benjamins.
- Vanhove, Martine. 2017. *Le bedja*. Leuven: Peeters.
- Vanhove, Martine & Mohamed-Tahir Hamid Ahmed. 2011. Le bedja. In Emilio Bonvini, Joëlle Bussutil & Alain Peyraube (eds.), *Dictionnaire des langues*, 362–368. Paris: PUF.
- Wedekind, Klaus. 2012. Sociolinguistic developments affecting beja dialects. In Matthias Brenzinger & Anne-Maria Fehn (eds.), *Proceedings of the 6th World Congress of African Linguistics, Cologne, 17–21 August 2009*, 623–633. Cologne: Rüdiger Köppe.
- Wohlgemuth, Jan. 2009. *A typology of verbal borrowings*. Berlin: Mouton de Gruyter.
- Zaborski, Andrzej. 1986. *The morphology of nominal plural in the cushitic languages*. Vienna: Afro-pub.
- Záhořík, Jan. 2007. The islamization of the beja until the 19th century. In Marc Seifert, Markus Egert, Fabian Heerbaart, Kathrin Kolossa, Mareike Limanski, Meikal Mumin, Peter André Rodekuhr, Susanne Rous, Sylvia Stankowski & Marilena Thanassoula (eds.), *Beiträge zur 1. Kölner Afrikawissenschaftlichen Nachwuchstagung (KANT I)*. Cologne: Institut für Afrikanistik, Universität zu Köln.
- Zibeli-Chen, Karola. 2014. Sprachen nubien in pharaonischer Zeit. *Lingua Aegyptia: Journal of Egyptian Language Studies* 22. 267–309.

Chapter 11

Berber

Lameen Souag

CNRS, Lacito

Arabic has influenced Berber at all levels – not just lexically, but phonologically, morphologically, and syntactically – to an extent varying from region to region. Arabic influence is especially prominent in smaller northern and eastern varieties, but is substantial even in the largest varieties; only in Tuareg has Arabic influence remained relatively limited. This situation is the result of a long history of large-scale asymmetrical bilingualism often accompanied by language shift.

1 Current state and contexts of use

1.1 Introduction

Berber, or Tamazight, is the indigenous language family of northwestern Africa, distributed discontinuously across an area ranging from western Egypt to the Atlantic, and from the Mediterranean to the Sahel. Its range has been expanding in the Sahel within recent times, as Tuareg speakers move southwards, but in the rest of this area, Berber has been present since before the classical period (Múrcia Sánchez 2010). Its current discontinuous distribution is largely the result of language shift to Arabic over the past millennium.

At present, the largest concentrations of Berber speakers are found in the highlands of Morocco (Tashelhiyt, Tamazight, Tarifiyt) and northeastern Algeria (Kabyle, Chaoui). Tuareg, in the central Sahara and Sahel, is more diffusely spread over a large but relatively sparsely populated zone. Across the rest of this vast area, Berber varieties constitute small islands – in several cases, single towns – in a sea of Arabic.

This simplistic map, however, necessarily leaves out the effects of mobility – not limited to the traditional practice of nomadism in the Sahara and transhumance in parts of the Atlas mountains. The rapid urbanisation of North Africa



Lameen Souag

over the past century has brought large numbers of Berber speakers into traditionally Arabic-speaking towns, occasionally even changing the town's dominant language. The conquests of the early colonial period created small Berber-speaking refugee communities in the Levant and Chad, while more recent emigration has led to the emergence of urban Berber communities in western Europe and even Quebec.

1.2 Sociolinguistic situation of Berber

In North Africa proper, the key context for the maintenance of Berber is the village. Informal norms requiring the use of Berber with one's relatives and fellow villagers, or within the village council, encourage its maintenance not only there but in cities as well, depending on the strength of emigrants' (often multigenerational) ties to their hometowns. In some areas, such as Igli in Algeria (Mouili 2013), the introduction of mass education in Arabic has disrupted these norms, encouraging parents to speak to their children in Arabic to improve their educational chances; in others, such as Siwa in Egypt (Serreli 2017), it has had far less impact. Beyond the village, in wider rural contexts such as markets, communication is either in Berber or in Arabic, depending on the region; where it is in Arabic, it creates a strong incentive for bilingualism independent of the state's influence. For centuries, Berber-speaking villages in largely Arabic-speaking areas have sporadically been shifting to Arabic, as in the Blida region of Algeria (El Arifi 2014); the opposite is also more rarely attested, as near Tizi-Ouzou in Algeria (Gautier 1913: 258).

In urban contexts, on the other hand, norms enforcing Berber have no public presence – quite the contrary. There one addresses a stranger in Arabic, or sometimes French, but rarely in Berber, except perhaps in a few Berber-majority cities such as Tizi-Ouzou (Tigziri 2008). Even within the family, Arabic takes on increasing importance; in a study of Kabyle Berbers living in Oran (Algeria), Ait Habbouche (2013: 79) found that 54% said they mostly spoke Arabic to their siblings, and 10% even with their grandparents. In the Sahel, Arabic is out of the picture, but there too family language choice is affected; 13% of the Berber speakers interviewed by Jolivet (2008: 146) in Niamey (Niger) reported speaking no Tamasheq at all with their families, using Hausa or, less frequently, Zarma instead.

Bilingualism is widespread but strongly asymmetrical. Almost all Berber speakers learn dialectal Arabic (as well as standard Arabic, taught at school), whereas Arabic speakers almost never learn Berber. There are exceptions: in some contexts, Arabic-speaking women who marry Berber-speaking men need to learn

Berber to speak with their in-laws (the author has witnessed several Kabyle examples), while Arabic speakers who settle in a strongly Berber-speaking town – and their children – sometimes end up learning Berber, as in Siwa (Egypt). Nevertheless, most Arabic speakers place little value on the language, and some openly denigrate it; in Bechar (Algeria), anyone expressing interest in Berber can expect frequently to hear the contemptuous saying *aš-šālḥa ma-hu klam wə-d-dhən ma hu l-idam* ‘Shilha (Berber) is no more speech than vegetable oil is animal fat’. To further complicate the situation, French remains an essential career skill (except in Libya and Egypt), since it is still the working language of many ministries and companies; in some middle-class families, it is the main home language spoken with children.

On paper, Berber (Tamazight) is now an official language of Morocco (since 2011) and Algeria (since 2016), while Tuareg (Tamasheq/Tamajeq) is a recognised national language of Mali and Niger. In practice, “official language” remains a misleading term. Official documents are rarely, if ever, provided in Berber, and there is no generalized right to communicate with the government in Berber. However, Berber is taught as a school subject in selected Algerian, Moroccan, and (since 2012 or so) Libyan schools, while some Malian and Nigerien ones even use it as a medium of education. It is also used in broadcast media, including some TV and radio channels. Both Morocco and Algeria have established language planning bodies to promote neologisms and encourage publishing, with a view towards standardisation. The latter poses difficult problems, given that each country includes major varieties which are not inherently mutually intelligible.

Berber varieties have been written since before the second century BC (Pichler 2007) – although the language of the earliest inscriptions is substantially different from modern Berber and decipherable only to a limited extent – and southern Morocco has left a substantial corpus of pre-colonial manuscripts (Boogert 1997); many other examples could be cited from long before people such as Mammeri (1976) attempted to make Berber a printed language. Nevertheless, writing seems to have had very little impact on the development of Berber as yet. Awareness of the existence of a Berber writing system – Tifinagh – is widespread, and often a matter of pride. However, most Berber speakers have never studied Berber, and do not habitually read or write in it in any script – with the increasingly important exception of social media and text messages, typically in Latin or Arabic script depending on the region. Efforts to create a standard literary Berber language have not so far been successful enough to exert a unifying influence on its dispersed varieties. In the North African context, this is often understood as implying that Berber is not a language at all – “language” (Arabic *luya*) being pop-

Lameen Souag

ularly understood in the region as “standardized written language”.

1.3 Demographic situation of Berber

No reliable recent estimate of the number of Berber speakers exists; relevant data is both scarce and hotly contested. The estimates brought together by Kossmann (2011: 1; 2013: 29–36) suggest a range of 30–40% for Morocco, 20–30% for Algeria, 8% for Niger, 7% for Mali, about 5% for Libya, and less than 1% for Tunisia, Egypt, and Mauritania. Selecting the midpoint of each range, and substituting in the mid-2017 populations of each of these countries (CIA 2017) would yield a total speaker population of about 25 million, 22 million of them divided almost evenly between Morocco and Algeria.

2 Contact languages and historical development

2.1 Across North Africa

Berber contact with Arabic began in the seventh century with the Islamic conquests. For several centuries, language shift seems to have been largely confined to major cities and their immediate surroundings, probably affecting Latin speakers more than Berber speakers. The invasion of the Banū Hilāl and Banū Sulaym in the mid-eleventh century is generally identified as the key turning point: it made Arabic a language of pastoralism, rapidly reshaping the linguistic landscape of Libya and southern Tunisia, then over the following centuries slowly transforming the High Plateau and the northern Sahara in general. This rural expansion further reinforced the role of Arabic as a lingua franca, while the recruitment of Arabic-speaking soldiers from pastoralist tribes encouraged its spread further west to the Moroccan Gharb.

The resulting linguistic divide between rural groups and towns remained a key theme of Maghrebi sociolinguistics until the twentieth century. In several cases, a town spoke a different language than its hinterland; in much of the Sahara, Berber-speaking oasis towns such as Ouargla or Igli formed linguistic islands in regions otherwise populated by Arabic speakers, and in the north, towns such as Bejaia or Cherchell constituted small Arabic-speaking communities surrounded by a sea of Berber-speaking villages. Even in larger cities such as Algiers or Marrakech, the dominance of Arabic was counterbalanced by substantial regular immigration from Berber-speaking regions further afield.

Today all Berber communities are more or less multilingual, usually in Arabic and often also in French; outside of the most remote areas, monolingual speakers

are quite difficult to find. Even in the nineteenth century, however, monolingual Berber speakers were considerably more numerous (Kossmann 2013: 41).

Alongside the coexistence of colloquial Maghrebi Arabic with Berber, Classical Arabic also had a role to play as the primary language of learning and in particular religious studies. Major Berber-speaking areas such as Kabylie (northern Algeria) and the Souss (southern Morocco) developed extensive systems of religious education, whose curricula consisted primarily of Arabic books (Boogert 1997; Mechehed 2007). The restriction of Classical Arabic to a limited range of contexts, and the relatively small proportion of the population pursuing higher education, gave it a comparatively small role in the contact situation; even in the lexicon, its influence is massively outweighed by that of colloquial Arabic, and it appears to have had no structural influence at all.

2.2 In Siwa

Examples of contact-induced change in this chapter are often drawn from Siwi, the Berber language of the oasis of Siwa in western Egypt. Sporadic long-distance contact with Arabic there presumably began in the seventh or eighth century with the Islamic conquests, and increased gradually as Cyrenaica and Lower Egypt became Arabic-speaking and as the trade routes linking Egypt to West Africa were re-established. During the eleventh century, the Banū Sulaym, speaking a Bedouin Arabic dialect, established themselves throughout Cyrenaica.

In the twelfth century, al-Idrīsī reports Arab settlement within Siwa itself, alongside the Berber population. Later geographers make no mention of an Arab community there, suggesting that these early immigrants were integrated into the Berber majority. Several core Arabic loans in Siwi, such as the negative copula *qačči* < *qaṭṭ šay?* and the noon prayer *luli* < *al-ṭūlē*, are totally absent from surrounding Arabic varieties today; such archaisms are likely to represent founder effects dating back to this period (Souag 2009).

The available data gives nothing close to an adequate picture of the linguistic environment of medieval Siwa. We may assume that, throughout these centuries, most Siwis – or at least the dominant families – would have spoken Berber as their first language, and more mobile ones – especially traders – would have learned Arabic (but whose Arabic?) as a second language. Alongside these, however, we must envision a fluctuating population of Arabic-speaking immigrants and West African slaves learning Berber as a second language. In such a situation, both Berber-dominant and Arabic-dominant speakers should be expected to play a part in bringing Arabic influences into Siwi.

The oasis was integrated into the Egyptian state by Muhammad Ali in 1820,

Lameen Souag

but large-scale state intervention in the linguistic environment of the oasis only took effect in the twentieth century; the first government school was built in 1928, and television was introduced in the 1980s. An equally important development during this period was the rise of labour migration, taking off in the 1960s as Siwi landowners recruited Upper Egyptian labourers, and Siwi young men found jobs in Libya's booming oil economy. It has then grown further since the 1980s with the rise of tourism and the growth of tertiary education. The effects of this integration into a national economy include a conspicuous generation gap in local second-language Arabic: older and less educated men speak a Bedouin-like dialect with *q > g, while younger and more educated ones speak a close approximation of Cairene Arabic.

3 Contact-induced changes in Berber

3.1 Introduction

As noted above, bilingualism in North Africa has been asymmetrical for many centuries, with Berbers much more likely to learn Arabic than vice versa. This suggests the plausible general assumption that the agents of contact-induced change were typically dominant in the (Berber) recipient language rather than in Arabic. However, closer examination of individual cases often reveals a less clear-cut situation; as seen above in §2.1, the history of Siwi suggests that Berber- and Arabic-dominant speakers both had a role to play, and *post facto* analysis of the language's structure seems to confirm this assumption. The loss of feminine plural agreement, for example (§3.3 below), can more easily be attributed to Arabic-dominant speakers adopting Berber than to Berber-dominant speakers. In the absence of clear documentary evidence, caution is therefore called for in the application of Van Coetsem's (1988; 2000) model to Berber.

3.2 Phonology

The influence of Arabic on Berber phonology is conspicuous; in general, every phoneme used in a given region's dialectal Arabic is found in nearby Berber varieties. Almost all Northern Berber varieties have adopted from Arabic at least the pharyngeals /ʕ/ and /ħ/, a series of voiceless emphatics: /s/, /h/, non-geminate /q/, and either /d/ or /t/. These phonemes presumably reached Berber through loanwords from Arabic, but have been extended to inherited vocabulary as well, through reinterpretation of emphatic spread or through their use in "expressive

formations” (Kossmann 2013: 199), e.g. Kabyle *θi-ħəðmər-θ* ‘breast of a small animal’ < *iðmar-ən* ‘breast’.

In Siwi (Souag 2013: 36–39; Souag & van Putten 2016), at least nine phonemes were clearly introduced from Arabic. The pharyngealised coronals /s/, /l/, /r/ and /d/ have no regular source in Berber, and occur in inherited vocabulary almost exclusively as a result of secondary emphasis spread (with the isolated exception of *dəs* ‘to laugh’). The order of borrowing appears to be *l, r > s > d*; in a few older loans, Arabic *s* is borrowed as *z* (e.g. *zəffar* ‘to whistle’ < *šaffar*), and in all but the most recent strata of loans, Arabic *d/ð* is borrowed as *t* (e.g. *a-ʃrit* ‘broad’ < *šarīd*). The pharyngeals /ħ/ and /ʕ/ (e.g. *ħabba* ‘a little’ < *ħabba* ‘a grain’, *ʃammi* ‘paternal uncle’ < *ʃamm-ī* ‘uncle-OBL.1SG’) likewise have no regular source in Berber, although 1SG -*y*- has become -*ʃ*- for some speakers (an irregular sound change specific to this morpheme). *ʃ* is lost in a number of older loans (e.g. *annaš* ‘bier’ < *an-našš*), but *ħ* is always retained as such rather than being dropped or adapted (unlike Tuareg, where it is typically adapted to *h*). This suggests that Siwi continued to adapt Arabic loans to its phonology by dropping *ʃ* up to some stage well after the beginning of significant borrowing from Arabic, but started accepting Arabic loans with *ħ* too early for any adapted to survive, implying an order of borrowing *ħ > ʕ*. Among the glottals, /h/ (e.g. *ddhan* ‘oil’ < *dihān* ‘oils’) appears in inherited vocabulary only in the distal demonstratives, where comparison to Berber languages that do have *h* suggests that it is excrescent, while /ʔ/ only rarely appears even in recent loanwords (e.g. *ʔəḡḡar* ‘to rent’ < *ʔaḡḡar*). The mid vowel /o/ has been integrated into Siwi phonology as a result of borrowing from Arabic; having been established as a phoneme, however, it went on to emerge by irregular change from original **u* in two inherited words (*allon* ‘window’, *agroz* ‘palm heart’), and from irregular simplification of **ayu* in some demonstratives (e.g. *wok* ‘this.SG.M’ < **wa yuṛ-ək* ‘this.SG.M at-2SG.M’). The interdental /θ/ and /ð/ have a more marginal status, but are used by some speakers even in morphologically well-integrated loans, e.g. *a-θqil* or *a-tqil* ‘heavy’ < *θaqīl*. Arabic influence may also be responsible for the treatment of [ʒ] and [dʒ] as free variants of the same phoneme /ǧ/ (Vycichl 2005), so that e.g. /taǧlaʃt/ ‘spider’ is variously realized as [tʰæʒlʰasʰt] ~ [tʰædʒlʰasʰt] (Naumann 2012: 152); other Berber languages with phonemic *ǰ* normally have [dʒ] as a conditioned allophone (e.g. when geminated) or as a cluster.

Arabic influence has also massively affected the frequency of some phonemes. /q/ and /ħ/ were marginal in Siwi before Arabic influence, while **e* had nearly disappeared due to regular sound changes, but all three are now quite frequent. Conversely, the influx of Arabic loans has helped make labiovelarised phonemes

Lameen Souag

such as g^w and q^w rare.

3.3 Morphology

Berber offers numerous examples of the borrowing of Arabic words together with their original Arabic inflectional morphology, a case of what [Kossmann \(2010\)](#) calls Parallel System Borrowing. This phenomenon is most prominent for nominal number marking, but sometimes attested in other contexts too.

In Berber, most nouns are consistently preceded by a prefix marking gender (masculine/feminine), number (singular/plural), and often case/state. Nouns borrowed from Arabic normally either get assigned a Berber prefix, or fill the prefix slot with an invariant reflex of the Arabic definite article: compare Figuig *a-gʃud* vs. Siwi *l̥a-gʃud* ‘young camel’ (< *qaʃūd*). The Berber plural marking system prior to Arabic influence was already rather complex, combining several different types of affixal marking with internal ablaut strategies; many Arabic loans are integrated into this system, e.g. Kabyle *a-bellar* ‘crystal’ > pl. *i-bellar-en* (< *billawr*), Siwi *a-kəddab* ‘liar’ > pl. *i-kəddab-ən* (< *kaððāb*). However, in most Berber varieties, Arabic loans have further complicated the system by frequently retaining their original plurals, e.g. Kabyle *l-kayəd* ‘paper’ > *le-kwayəd* (< *kāyid*), Siwi *əl-gənʃud* ‘hedgehog’ > pl. *l̥a-gnaʃid* (< *qunʃud*). (The difference correlates fairly well with the choice in the singular between a Berber prefix and an Arabic article, but not perfectly; contrast e.g. Siwi *a-ʃruḥ* ‘chick, bastard’ < *ʃarḥ*, which takes the Arabic-style plural *l̥a-ʃraḥ*.) Berber has no inherited system of dual marking, instead using analytic strategies. Nevertheless, for a limited number of measure words, duals too are borrowed, e.g. Kabyle *yum-ayen* ‘two days’ < *yawm-aʃyn* (although ‘day’ remains *ass*!), Siwi *s-sən-t* ‘year’ > *sən-t-en* ‘two years’ < *san-at-aʃyn*. Arabic number morphology may sporadically spread to inherited terms as well, e.g. Kabyle *berdayen* ‘twice’ < *a-brid* ‘road, time’, Siwi *l̥a-gʷrazən* ‘dogs’ < *a-gʷarzni* ‘dog’ ([Souag 2013](#)).

Whereas nouns are often borrowed together with their original inflectional morphology, verbs almost never are. The only attested exception is Ghomara, a heavily mixed variety of northern Morocco. In Ghomara, many (but not all) verbs borrowed from Arabic are systematically conjugated in Arabic in otherwise monolingual utterances, a phenomenon which seems to have remained stable over at least a century: thus ‘I woke up’ is consistently *faq-aḥ*, but ‘I fished’ is equally consistently *ʃʃað-iθ* ([Mourigh 2016](#): 6, 137, 165). However, the borrowing of Arabic participles to express progressive aspect is also attested in Zuwara, if only for the two verbs of motion *mašəy* ‘going’ (pl. *mašy-in*) and *žay* ‘coming’ (pl. *žayy-in*), contrasting with inherited *fəl* ‘go’, *asəd* ‘come’ ([Kossmann 2013](#): 284–

285).

Prepositions are less frequently borrowed; in some cases where they are borrowed, however – including Igli *mənyir-* ‘except’, Ghomara *bin* ‘between’ (Kossmann 2013: 293) – they too occasionally retain Arabic pronominal markers, e.g. Siwi *msabb-ha* ‘for her’ < *min sababi-hā* ‘from reason.OBL-OBL.3SG.F’ (Souag 2013: 48). In Awjila, more unusually, two inherited prepositions somewhat variably take Arabic pronominal markers, e.g. *dit-ha* ‘in front of her’ (van Putten 2014: 113).

A rarer but more spectacular example of morpheme borrowing is the borrowing of productive templates from Arabic. Such cases include the elative template əCCəC in Siwi, used to form the comparative degree of trilateral adjectives irrespective of etymology – thus *əmləl* ‘whiter’ < *a-məllal* alongside *ətwəl* ‘taller’ < *a-ṭwil* < Arabic *ṭawīl* (Souag 2009) – and the diminutive template CCiCəC in Ghomara (Mourigh 2016), e.g. *aṣwiyyər* ‘little root’ < *aṣar* alongside *ləmwiyyəs* ‘little knife’ < *l-mus* < Arabic *al-mūsā* ‘razor’ (gemination of *y* is automatic in the environment i_V). As the latter example illustrates, borrowed derivational morphology sometimes becomes productive.

The effects of Arabic on Berber morphology are by no means limited to the borrowing of morphemes. There is reason to suspect Arabic influence of having played a role in processes of simplification attested mainly in peripheral varieties, such as the loss of case marking in many areas. In Siwi, where Arabic influence appears on independent grounds to be unusually high, the verbal system shows a number of apparent simplifications targeting categories absent in sedentary Arabic varieties: the loss of distinct negative stems, the near-complete merger of perfective with aorist, the fixed postverbal position of object clitics, and so on. It is tempting to explain such losses as arising from imperfect acquisition of Siwi by Arabic speakers.

Structural calquing in morphology is also sporadically attested. Siwi has lost distinct feminine plural agreement on verbs, pronouns, and demonstratives, extending the inherited masculine plural forms to cover plural agreement irrespective of gender. Within Berber, this is unprecedented; plural gender agreement is extremely well conserved across the family. However, it perfectly replicates the usual sedentary Arabic system found in Egypt and far beyond.

3.4 Syntax

Syntactic influence is often difficult to identify positively. Nevertheless, Berber offers a number of examples, of which relative clause formation is one of the clearest (Souag 2013: 151–156; Kossmann 2013: 369–407). Relative clauses in Berber

Lameen Souag

are normally handled with a gap strategy combined with fronting of any stranded prepositions, as in (1).

- (1) Awjila (Paradisi 1961: 79)
 ərrafəqa-nnəs wi ižin-an-a nettin id-sin ksum
 friend.PL-GEN.3SG REL.PL.M divide-3PL.M-PRF 3SG.M with-OBL.3PL.M meat
 ‘his friends with whom he divided the meat’

In subject relativisation, a special form of the verb not agreeing in person (the so-called “participle”) is used, as in (2); such a form is securely reconstructible for proto-Berber (Kossmann 2003).

- (2) Awjila (Paradisi 1960: 162)
 amədən wa tarəv-ən nettin fayyan
 man REL.SG.M write.IPFV-PTCP 3SG.M ill
 ‘The man who is writing is ill.’

In several smaller easterly varieties apart from Awjila, however, both of these traits have been lost. The strategy found in varieties such as Siwi – resumptive weak (affixal) pronouns throughout, and regular finite agreement for subject relativisation – perfectly parallels Arabic:

- (3) Siwi (Souag 2013: 151–152)
 tálti tən dəzz-γ-as ġğəwab
 woman REL.SG.F send-1SG-DAT.3SG letter
 ‘the woman to whom I sent the letter’
- (4) Siwi (field data)
 ágg^wid wənn i-ḡəṃṃar iməğran
 man REL.SG.M 3SG.M-make.IPFV sickle.PL
 ‘the man who makes sickles’

In the case of verbal negation, an originally syntactic calque has often been morphologised in parallel in Arabic and Berber. A number of varieties – especially the widespread Zenati subgroup of Berber, ranging from eastern Morocco to northern Libya – have developed a postverbal negative clitic *š* / *ša* from **kāra* ‘thing’, apparently a calque on Arabic *š* / *ši* from *šay?*; however, some instead use the direct borrowings *ši* or *šay* (Lucas 2007; Kossmann 2013: 332–334).

3.5 Lexicon

Lexical borrowing from Arabic is pervasive in Berber. Out of 41 languages around the world compared in the Loanword Typology Project (Tadmor 2009), Tarifiyt Berber was second only to (Selice) Romani in the percentage of loanwords – more than half (51.7%) of the concepts compared. More than 90% of loanwords examined in Tarifiyt were from Arabic, almost all from dialectal Maghrebi Arabic. There is little reason to suppose that Tarifiyt is exceptional in this respect among Northern Berber languages; to the contrary, Kossmann (2013: 110) finds its rate of basic vocabulary borrowing to be typical of Northern Berber, whereas Siwi and Ghomara go much higher. The rate of borrowing from Arabic, however, is considerably lower further south and west; on a 200-word list of basic vocabulary, Chaker (1984: 225–226) finds 38% Arabic loans in Kabyle (north-central Algeria) vs. 25% in Tashelhiyt (southern Morocco) and only 5% in Tahaggart Tuareg (southern Algeria).

This borrowing is pervasive across the languages concerned, rather than being restricted to particular domains. Every semantic field examined for Tarifiyt, including body parts, contained at least 20% loanwords, and verbs or adjectives were about as frequently borrowed as nouns were (Kossmann 2009). Numerals stand out for particularly massive borrowing; most Northern Berber varieties have borrowed all numerals from Arabic above a number ranging from ‘one’ to ‘three’ (Souag 2007).

The effects of this borrowing on the structure of the lexicon remain insufficiently investigated, but appear prominent in such domains as kinship terminology. Throughout Northern Berber, a basic distinction between paternal kin and maternal kin is expressed primarily with Arabic loanwords (*ṣammi* ‘paternal uncle’ vs. *ḥali* ‘maternal uncle’ etc.), whereas in Tuareg that distinction is not strongly lexicalised. Nevertheless, borrowing does not automatically entail lexical restructuring; Tashelhiyt, for example, kept its vigesimal system even after borrowing the Arabic word for ‘twenty’ (*ṣṣrin*), cf. Ameur (2008: 77).

The borrowing of analysable multi-word phrases – above all, numerals followed by nouns – stands out as a rather common outcome of Berber contact with Arabic. Usually this is limited to the borrowing of numerals in combination with a limited set of measure words, such as ‘day’; thus in Siwi we find forms like *sbaṣ-t iyyam* ‘seven days’ rather than the expected regular formation **səbṣa n nnhar-at* (Souag 2013: 114). In Beni Snous (western Algeria), the phenomenon seems to have gone rather further: Destaing (1907: 212) reports that numerals above ‘ten’ systematically select for Arabic nouns. Souag & Kherbache (2016), however, explain this as a codeswitching effect, rather than a true case of one

Lameen Souag

language's grammar requiring shifts into another.

4 Conclusion

The influence of Arabic on Berber has come to be better understood over the past couple of decades, but much remains to be done. Synchronically, Berber–Arabic codeswitching remains virtually unresearched; rare exceptions include Hamza (2007) and Kossmann (2012). Sociolinguistic methods could help us better understand the gradual integration of new Arabic loanwords; the early efforts of Brahimi (2000) have hardly been followed up on. Diachronically, it remains necessary to move beyond the mere identification of loanwords and contact effects towards a chronological ordering of different strata, an approach explored for some peripheral varieties by Souag (2009) and van Putten & Benkato (2017). While linguists are belatedly beginning to take advantage of earlier manuscript data to understand the history of Berber (Boogert 1997; Boogert 1998; Brugnatelli 2011; Meouak 2015), this data has not yet been used in any systematic way to help date the effects of contact at different periods. For many smaller varieties, especially in the Sahara, basic documentation and description are still necessary before the influence of Arabic can be explored. The unprecedented degree of Arabic influence revealed in Ghomara by recent work (Mourigh 2016), extending to the borrowing of full verb paradigms, suggests that such descriptive work may yet yield dividends in the study of contact.

Despite all these gaps, the work done so far is more than sufficient to establish a general picture of Arabic influence on Berber. Throughout Northern Berber, Arabic influence on the lexicon is substantial and pervasive, bringing with it significant effects on phonology and morphology. Structural effects of Arabic on morphology, and Arabic influence on Berber syntax, are less conspicuous but nevertheless important, especially in smaller varieties such as Siwi. Looking at these results through Van Coetsem's framework, this suggests that RL-dominant speakers have had an especially dominant role in Arabic–Berber contact in larger varieties, whereas SL-dominant speakers' role is more visible in smaller varieties. However, this *a priori* conclusion should be tested against directly attested historical data wherever possible.

Further reading

The key reference for Arabic influence on Northern Berber is Kossmann (2012), frequently cited above; this covers all levels of influence including the lexicon,

phonology, nominal and verbal morphology, borrowing of morphological categories, and syntax.

The most extensive in-depth study of Arabic influence on a specific Berber variety is [Souag \(2013\)](#), effectively a contact-focused grammatical sketch of Siwi Berber.

[Mourigh \(2016\)](#) is a thorough synchronic description of by far the most strongly Arabic-influenced Berber variety, Ghomara, giving a uniquely clear picture of just how far the process can go without resulting in language shift.

Acknowledgements

The author thanks his consultants in Siwa, especially the late Sherif Bougdoura, for their help with studying Siwi.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	OBL	oblique
DAT	dative	PL/pl.	plural
F	feminine	PRF	perfect
GEN	genitive	PTCP	participle
IPFV	imperfective	SG	singular
M	masculine		

References

Ait Habbouche, Khadidja. 2013. *Language maintenance and language shift among kabyle speakers in Arabic speaking communities: The case of Oran*. Oran: University of Oran master’s dissertation.

Ameur, Meftaha. 2008. Les contacts arabe-amazighe: le cas des noms de nombre. In Mena Lafkioui & Vermondo Brugnatelli (eds.), *Berber in contact: Linguistic and sociolinguistic perspectives*, 63–80. Cologne: Rüdiger Köppe.

Boogert, Nico van den. 1997. *The Berber literary tradition of the Sous, with an edition and translation of “The Ocean of Tears” by Muḥammad Awzal (d. 1749)*. Leiden: Nederlands Instituut voor het Nabije Osten.

Boogert, Nico van den. 1998. “La révélation des énigmes”: *Lexiques arabo-berbères des XVIIe et XVIIIe siècles*. Aix-en-Provence: Institut de recherches et d’études sur le monde arabe et musulman (IREMAM).

Lameen Souag

- Brahimi, Fadila. 2000. Loanwords in Algerian Berber. In Jonathan Owens (ed.), *Arabic as a Minority Language*, 371–382. Berlin: Mouton de Gruyter.
- Brugnatelli, Vermondo. 2011. Some grammatical features of ancient Eastern berber (the language of the mudawwana). In Luca Busetto (ed.), *Studies on language and African linguistics in honour of Marcello Lamberti*, 29–40. Milano: Qu.A.S.A.R.
- Chaker, Salem. 1984. *Textes en linguistique berbère: Introduction au domaine berbère*. Paris: Editions du Centre national de la recherche scientifique.
- CIA. 2017. *The world factbook 2017*. Washington, D. C.: Central Intelligence Agency.
- Destaing, Edmond. 1907. *Etude sur le dialecte berbère des Beni-Snous*. Paris: Ernest Leroux.
- El Arifi, Samir. 2014. *Tachelhit de l'Atlas blidéen*. Fourth edition. Algiers: Haut Commissariat à l'Amazighité.
- Gautier, E.-F. 1913. Répartition de la langue berbère en algérie. *Annales de géographie* 22(123). 255–266.
- Hamza, Belgacem. 2007. *Berber ethnicity and language shift in Tunisia*. Brighton: University of Sussex doctoral dissertation.
- Jolivet, Remi. 2008. La fidélité au tamashek des berbérophones dans le contexte multilingue du Niger. In Mena Lafkioui & Vermondo Brugnatelli (eds.), *Berber in contact: Linguistic and sociolinguistic perspectives*, 139–149. Cologne: Rüdiger Köppe.
- Kossmann, Maarten. 2003. The origin of the berber “participle”. In M. Lionel Bender, Gábor Takács & David L. Appleyard (eds.), *Selected comparative–historical Afrasian linguistics in the memory of Igor M. Diakonoff*, 27–40. Munich: LINCOM Europa.
- Kossmann, Maarten. 2009. Loanwords in tarifiyt. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 191–214. Berlin: De Gruyter.
- Kossmann, Maarten. 2010. Parallel system borrowing: parallel morphological systems due to the borrowing of paradigms. *Diachronica* 27(3). 459–487.
- Kossmann, Maarten. 2011. *A grammatical sketch of Ayer Tuareg (Niger)*. Cologne: Rüdiger Köppe.
- Kossmann, Maarten. 2012. Berber–arabic code-switching in Imouzzar du Kandar. *STUF – Language Typology and Universals. Sprachtypologie und Universalienforschung* 65(4). 369–382.
- Kossmann, Maarten. 2013. *The Arabic influence on Northern Berber*. Leiden: Brill.

- Lucas, Christopher. 2007. Jespersen's cycle in Arabic and berber. *Transactions of the Philological Society* 105(3). 398–431.
- Mammeri, Mouloud. 1976. *Tajerrumt n tmazight (tantala taqbaylit)*. Paris: Maspéro.
- Mechehed, Djamel-Eddine. 2007. L'organisation des notices de catalogage des manuscrits arabes ou berbères: cas de la collection Ulahbib, Kabylie. In *Les manuscrits berbères au Maghreb et dans les collections européennes*, 129–138. La Fresquièrre: Atelier Perrousseaux.
- Meouak, Mohamed. 2015. *La langue berbère au Maghreb médiéval*. Leiden: Brill.
- Mouili, Fatiha. 2013. The Berber language of Igli: language towards extinction. *Humanities and Social Sciences Review* 2(2). 563–580.
- Mourigh, Khalid. 2016. *A grammar of Ghomara Berber (north-west Morocco)*. Cologne: Rüdiger Köppe.
- Múrcia Sánchez, Carles. 2010. *La llengua amaziga a l'antiguitat a partir de les fonts gregues i llatines*. Barcelona: Universitat de Barcelona.
- Naumann, Christfried. 2012. *Acoustically based phonemics of siwi (berber)*. Cologne: Rüdiger Köppe.
- Paradisi, Umberto. 1960. Il berbero di augila: materiale lessicale. *Rivista degli Studi Orientali* 35. 157–177.
- Paradisi, Umberto. 1961. Testi berberi di augila (cirenaïca). *Annali dell'Istituto Orientale di Napoli* 10. 79–91.
- Pichler, Werner. 2007. *Origin and development of the Libyco-Berber script*. Cologne: Rüdiger Köppe.
- Serrelli, Valentina. 2017. Language and identity in Siwa Oasis: Indexing belonging, localness and authenticity in a small minority community. In Reem Bassiouney (ed.), *Identity and dialect performance: A study of communities and dialects*, 226–242. London: Routledge.
- Souag, Lameen. 2007. The typology of number borrowing in berber. In Naomi Hilton, Rachel Arscott, Katherine Barden, Sheena Shah & Meg Zellers (eds.), *Proceedings for the Fifth Cambridge Postgraduate Conference in Linguistics*, 237–244. Cambridge: Cambridge Institute of Language Research.
- Souag, Lameen. 2009. Siwa and its significance for Arabic dialectology. *Zeitschrift für Arabische Linguistik* 51. 51–75.
- Souag, Lameen. 2013. *Berber and arabic in siwa (egypt): a study in linguistic contact*. Cologne: Rüdiger Köppe.
- Souag, Lameen & Fatma Kherbache. 2016. Syntactically conditioned code-switching? The syntax of numerals in Beni-Snous Berber. *International Journal of Bilingualism* 20(2). 97–115.

Lameen Souag

- Souag, Lameen & Marijn van Putten. 2016. The origin of mid vowels in Siwi. *Studies in African Linguistics* 45(1–2). 189–208.
- Tadmor, Uri. 2009. Loanwords in the world's languages: Findings and results. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 55–75. Berlin: De Gruyter.
- Tigziri, Nora. 2008. Le kabyle au contact des langues en présence en algérie: entre codeswitching et parler hybride? In Mena Lafkioui & Vermondo Brugnatelli (eds.), *Berber in contact: Linguistic and sociolinguistic perspectives*, 175–185. Cologne: Rüdiger Köppe.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- van Putten, Marijn. 2014. *A grammar of awjila berber (libya), based on umberto paradisi's material*. Cologne: Rüdiger Köppe.
- van Putten, Marijn & Adam Benkato. 2017. The Arabic strata in Awjila Berber. In Ahmad Al-Jallad (ed.), *Arabic in context: Celebrating 400 years of Arabic at Leiden University*, 476–502. Leiden: Brill.
- Vycichl, Werner. 2005. Sprachskizzen: Die berberischen sprachinseln in tunesien. In Dymitr Ibriszimow & Maarten Kossmann (eds.), *Berberstudien & A sketch of Siwi Berber (Egypt)*, 136–147. Cologne: Rüdiger Köppe.

Chapter 12

Modern South Arabian Languages

Simone Bettega

Università degli Studi di Torino

Fabio Gasparini

Università degli Studi di Napoli “L’Orientale”

In the course of this chapter we will discuss what is known about the effects that contact with Arabic has had on the Modern South Arabian Languages of Oman and Yemen. Documentation concerning these languages is not abundant, and even more limited is our knowledge of the history of their interaction with Arabic. By integrating the existing bibliography with as yet unpublished fieldwork materials, we will try to provide as complete a picture of the situation as possible, also discussing the current linguistic and sociolinguistic landscape of Dhofar and Eastern Yemen.

1 History of contact between Arabic and the Modern South Arabian Languages

Much to the frustration of modern scholars of Semitic, the history of the Modern South Arabian Languages (henceforth MSAL) remains largely unknown.¹ To this day, no written attestation of these varieties has been discovered, and it seems safe to assume that they have remained exclusively spoken languages throughout all of their history. Since European researchers became aware of their existence in the first half of the nineteenth century (Wellsted 1837), and until very recently, the MSAL were thought by many to be the descendants of the old (epigraphic)

¹§1 was authored by Simone Bettega, while §2 was authored by Fabio Gasparini. §3 and §4 are the result of the conjoined efforts of both authors. In particular, Gasparini was responsible for analyzing most of the primary sources and raw linguistic data, while Bettega worked more extensively on the existing bibliography.



Simone Bettega & Fabio Gasparini

South Arabian languages (Rubin 2014: 16). This assumption has been conclusively disproven by Porkhomovsky's 1997 article, which also contributed in a significantly to the re-shaping of the proposed model for the Semitic family tree. This modified version of the family tree (which finds further support in the recent works of Kogan 2015 and Edzard 2017) sets the MSAL apart as an independent branch of the West Semitic subgroup, one whose origins are therefore of considerable antiquity. This brings us to the question of when it was that the MSAL (or their forebears) first came into contact with Arabic. This might have happened at any time since Arabic-speaking people started to penetrate into southern Arabia, a process that – as we know from historical records – began in the second half of the first millennium B.C. (Robin 1991; Hoyland 2001: 47–48). Roughly one thousand years later, almost the whole population of central and northern Yemen was speaking Arabic, and possibly a considerable portion of the southern population as well (Beeston 1981: 184; Zammit 2011: 295). It is therefore possible that Arabic and the MSAL have been in contact for quite some time, and it seems likely that the intensity and effects of such contact grew stronger after the advent of Islam (Lonnet 2011: 247). It is also possible, as some scholars have written, that the MSAL “represent isolated forms that were never touched by Arabic influence until the modern period” (Versteegh 2014: 127). Admittedly, evidence to support either one of these hypotheses is scarce, and at present it is probably safer to say that our knowledge of the history of contact between Arabic and the MSAL before the twentieth century is fragmentary at best. This is why studies on the outcomes of such contact are of particular interest, since they could help to shed light on parts of that history. This is also why, in the course of this chapter, we will refrain from addressing the question of how contact with the MSAL affected the varieties of Arabic spoken in Oman and Yemen, and focus solely on the influence of Arabic on the MSAL. Although there is plenty of evidence that South Arabian exerted a powerful influence on the Arabic of the area (see for instance Retsö 2000 and Watson 2018),² it is often difficult to assess whether this influence is the result of contact with forms of Ancient South Arabian or more recent interaction with the MSAL. Such a discussion, also because of space constraints, is beyond the scope of the present article.

As far as the interaction between Arabic and the MSAL in the twentieth and 21st centuries is concerned, Morris (2017: 25) provides a good overview of the multilingual environment in which the MSAL were and are spoken:

²To the point that so-called mixed varieties are reported to exist, whose exact linguistic nature seems difficult to pinpoint. See Watson et al. (2006) and Watson (2011) for discussion.

Speakers of [a Modern South Arabian] language always had to deal with speakers of other MSAL, as well as with speakers of various dialects of Arabic. The Baṭāḥirah, for instance, did nearly all their trade with boats from Ṣūr and other Arabic-speaking ports; they lived and worked with the Arabic-speaking Janaba, while being in contact with speakers of Ḥarāsīs and Mahra. The Ḥarāsīs interacted with the Arabic speakers surrounding their Jiddat al-Ḥarāsīs homeland, traded in the Arabic-speaking markets of the north, and in the summer months went to work at the northern date harvest. Mehri speakers lived beside and traded with Arabic-speaking Kathīri tribesmen in the Nejd region, Ṣheret speakers in the mountains, and Arabic speakers in the coastal market towns of Dhofar. Ṣheret speakers interacted with the Mahra, some of whom settled among them, and with Arabic-speaking peoples of the coast as well as the desert interior [...] There was marriage between Arabic-speaking men of the coastal towns and MSAL-speaking women of the interior, and over time, families of Mehri and Ṣheret speakers settled in or near the towns, with the result that even more Arabic speakers became familiar with these languages.

2 Current state of contact between Arabic and the Modern South Arabian Languages

Today, six Modern South Arabian Languages exist, spoken by around 200,000 people in Eastern Yemen (including the island of Soqatra) and Western Oman. These six languages are: Mehri, Hobyōt, Ḥarsūsi, Baṭḥari, Ṣheret/Jibbālī and Soqōṭri. They are all to be regarded as endangered varieties, though the individual degree of endangerment varies remarkably. No exact census concerning the number of speakers is currently available (Simeone-Senelle 2011: 1075), but we know that Mehri is the most spoken language, with an estimate of around 100,000 speakers. It is followed by Soqōṭri (about 50,000 speakers), Ṣheret (25,000), Ḥarsūsi (a few hundred), Hobyōt (a few hundred) and Baṭḥari (less than 20 speakers). The main causes of endangerment are reckoned to be shift to Arabic and disappearance of traditional local lifestyles. In addition, the current political situation in Yemen is having effects on the linguistic landscape of the region which are difficult to document or foresee: the area is currently inaccessible to researchers, and there is no way to know how the conflict will affect the local communities.

As far as Oman is concerned, the city of Salalah undoubtedly represents the major locus of contact between Arabic and the MSAL. The rapid growth the city has witnessed in recent years, and the improved possibilities of economic de-

Simone Bettega & Fabio Gasparini

velopment that came with it, have led many Šheret-speaking breeders from the nearby mountains to settle in the city or its immediate surroundings, where they now employ Arabic on a daily basis as a consequence of mass education and media neglecting other local languages. This has led to a split, in the speakers' perception, between "proper" Šheret, spoken in the mountains, and the "city Šheret" of Salalah, often regarded as a sort of "broken" variety of the language in which, among other things, code-switching with Arabic is extremely frequent. Unfortunately, data on this subject are virtually non-existent, given the extreme difficulty of documenting such an episodic phenomenon (aggravated by the speakers' understandable reluctance to having their imperfect language proficiency evaluated and recorded).

Even outside the urban centers, however, contact with Arabic is on the rise. Even the most isolated variety, Soqotri, is apparently undergoing rapid change under the influence of Arabic: the existence of a koineised variety of Soqotri, heavily influenced by Arabic, has been recently reported in Ḥadibo (Morris 2017: 27). This is not to say, of course, that all MSAL are being affected to the same degree: Watson (2012: 3), for instance, notes how "Mahriyōt [the Eastern Yemeni variety of Mehri] [...] exhibits structures unattested in Mehreyyet [Mehri Omani variety] [...] and shows greater Arabic influence both in terms of the number of Arabic terms used, and the length and frequency of Arabic phrases within texts." However, no MSAL seems at present to be exempt from the effects of contact.

The case of Baḥari, which, as we have seen, is the most severely endangered of all the MSAL, exemplifies well the processes of morphological loss and erosion that a language undergoes in the final stages of endangerment. Morris (2017: 12) reports how already in the 1970s Baḥari seemed to display many of the signs of a moribund language. In recent times

[t]he younger generations showed little interest in their former language; they were eager to embrace Arabic and to feel themselves part of the wider Arabic Islamic community; and they were proud to call themselves 'ʿarab', with all that word's overtones of Bedouin ancestry and code of honour (Morris 2017: 11).

In the following sections we will discuss several types of contact-induced changes in the MSAL: although we will use material taken from all varieties, Baḥari will be in particular focus due to its singular status.

3 Contact-induced changes in the MSAL

As already noted, in the course of this chapter we will focus solely on the effects that contact with Arabic has had on the various MSAL. Therefore, Arabic will always be the source language of all the transfer phenomena considered in the next pages, while the recipient language will be, depending on the different examples, one or the other of the six MSAL. Obviously, this poses the question of who the agents of change are and have been in the case of these particular phenomena, and what type(s) of transfer are we confronted with (cf. [Van Coetsem 1988; 2000; Winford 2005](#)). According to the overview of the MSAL's sociolinguistic status presented above, it should be clear by now that, while the two cases are extremely common of (a) mono- or multilingual MSAL speakers who acquire Arabic as an L2 and (b) bilingual MSAL-Arabic speakers, the opposite is not true (that is, monolingual Arabic speakers who come to acquire one or more MSAL as L2s later in life). In other words, all the transfer phenomena we will be considering in the next paragraphs are either instances of borrowing (brought about by speakers who are dominant in one or more MSAL) or convergence (brought about by speakers who are native speakers of Arabic and at least one MSAL; see [Lucas 2015](#) for a definition of convergence).

3.1 Phonology

3.1.1 Phonetic adaptation of loanwords

As illustrated in §3.4, lexical borrowings from Arabic are extremely common in the MSAL. As [Morris \(2017: 13\)](#) remarks, such loanwords are often altered in order for them to acquire a “South Arabian flavour”, so to speak. The phenomenon is not one of simple adaptation dictated by difficulty of articulation, since the sounds that are replaced are present in the phonological inventory of the MSAL. In fact, the opposite appears to be true, these sounds normally being replaced by others which are typical of South Arabian but absent in Arabic. For Baṭḥari, [Morris](#) gives the example of Arabic pharyngealised dental fricative /ð̣/ (IPA [ð̣ʰ]) being replaced by the pharyngealised alveolar lateral fricative /ṣ̣/ (mostly realised as IPA [ɬ̣ʰ], see §3.1.4), as in *raṣṣ* ‘bruise’ (from Janaybi Arabic *raḏḏ*), or Arabic /š/ (IPA [ʃ]) being replaced by /ṣ/, as in *men šān-k*, ‘for you, for your sake’, in place of *men šān-k*, *šarray* ‘buyer’ for *šarray*, or *šāmāl* ‘inland, north’ for *šāmāl* (while Baṭḥari *šēmāl(i)* is normally used to refer to the left hand only).

Lexical borrowing can also be the cause of variation in the realisation of certain sounds, as is the case with the phonemes /g/ and /y/ (IPA [g] and [j] respectively),

Simone Bettega & Fabio Gasparini

which represent different reflexes of Proto-Semitic *g in different Omani Arabic dialects. It is possible to find traces of this variation in those MSAL that are in contact with more than one variety of Arabic, as is the case with Ḥarsusi: see for instance *fagr* and *fayr*, both meaning ‘dawn’, or the opposition between *yann* ‘madness’ and *genni* ‘jinni’, both from the same etymological root (Lonnet 2011: 299).

3.1.2 Affrication of /k/ > [tʃ]

It can also be the case that some phonetic processes regularly taking place in the local Arabic varieties but otherwise unknown to MSAL phonology are transferred to original MSAL vocabulary. This is what happens in Baḥari, where some speakers may show an affricate realisation of the voiceless occlusive [k] > [tʃ], which resembles the Janaybi Arabic realisation of the phoneme /k/ (whose complementary distribution with the voiceless plosive realisation [k] is still unclear). For example, some speakers regularly produce /yənkaʃ/ ‘come.3SG.M.SBJV’ as [jən'tʃaʃ] instead of [jən'kaʃ].

3.1.3 Stress

The structural similarity of Arabic and the MSAL can sometimes cause stress patterns which are typical of the former to be applied to the latter, as is the case with ‘she began’: Soqotri *bédʔəh*, (local) Arabic *bədaʔat*, Soqotri with an Arabic stress *bədaʔəh* (Lonnet 2011: 299).

3.1.4 Realisation of emphatics

This is a topic that has attracted the attention of several scholars since the publication of Johnstone’s 1975 article on the subject, because of the realisation of the so-called Semitic “emphatics” as glottalised consonants. Glottalisation is a secondary articulatory process in which narrowing (creaky voice) or closure (ejective realisation) of the glottis takes place: the action of the larynx compresses the air in the vocal tract which, once released, produces a greater amplitude in the stop burst (Ladefoged & Maddieson 1996: 78).

Lonnet (2011: 299) notes a tendency for speakers of various MSAL to replace the ejective articulation of certain consonants (especially fricatives, see Ridouane & Gendrot 2017) with a pharyngealised realisation, typical of Arabic emphatics. Pharyngealisation is a kind of secondary articulation involving a constriction of the pharynx usually realised through tongue root retraction, resulting in a backed

realisation (Ladefoged & Maddieson 1996: 365), and it is a well-documented process across Semitic languages. Naumkin & Porkhomovsky (1981) note for Soqotri an ongoing process of transition from a glottalised to a pharyngealised realisation of emphatics, with only stops being realised as fully glottalised items. Work from Watson & Bellem (2010; 2011) and Watson & Heselwood (2016) show the co-occurrence of pharyngealisation and glottalisation in relation to pre-pausal phenomena in Şanġāni Arabic, Mahryōt and Mehreyyet (respectively the westernmost Yemeni and Omani varieties of Mehri). Dufour (2016: 22) states that “le caractère éjectif des phonèmes emphatiques ne fait aucun doute, en jibbali comme en mehri” (“the nature of the emphatic phonemes is undoubtedly ejective, in Jibbali as much as in Mehri”).³ Finally, in Baṭṭhari only /k/ is realised as a fully ejective consonant [kʰ]. /t/ and the fricative emphatics, on the other hand, are described as mainly pharyngealised (and partially voiced, in the case of fricatives; Gasparini 2017).

Unfortunately, since there is no thorough phonetic description of any MSAL that predates the 1970s, it is impossible to ascertain whether these realisations (which, again, range from fully glottalised to fully pharyngealised) are the result of the influence of Arabic, or have arisen as the consequence of internal and typologically predictable developments. It is likely, though, that bilingualism and constant contact with Arabic have at least favoured this phonetic change. Evidence in support of this view may come from the fact that speakers who are poorly proficient in Arabic and live in rural and more isolated areas are more likely to preserve a glottalised realisation of the emphatics (as emerges from direct fieldwork observations).

3.2 Morphology

3.2.1 Nominal morphology

Morphological patterns which are typical of Arabic can enter the language through borrowing, as is the case of the passive participle of simple verbs, which is *mVC-CūC* in Arabic and *mVCCīC* in MSAL. Soqotri *maḥlōk*, for instance, is clearly derived from Arabic *maḥlūq* ‘human being’ (lit. ‘created’), while this is not the case for Ḥarsusi *mḥaliḳ* (Lonnet 2011: 299). Also, in the realm of verbal derivational morphology, certain phenomena can be introduced into the recipient language through lexical borrowing: this is the case with gemination and prefixation of *t*- in Ḥarsusi, as in the participle *məṭṭaḥfi* ‘barefoot’ (from Omani Arabic *miṭṭaḥfi*; Lonnet 2011).

³ Authors’ translation.

Simone Bettega & Fabio Gasparini

In general, Arabic loanwords are normally well integrated in MSAL morphology, probably because of the high degree of structural similarity that exists between these languages. One example, reported by [Lonnet \(2011\)](#), is that of *bəkerēt* ‘cow’, a fully integrated loan from Arabic used in Ḥarsusi and Western Yemeni Mehri, which possesses its own plural and diminutive form (*bəḵār* and *bəḵārēnōt*, respectively).

Arabic loans in several MSAL stand out because of their characteristic feminine ending in *-(h)* instead of *-(V)t*, as in Ṣhehri *saṣah* ‘watch’ and *torah* ‘revolution’ (but consider the more adapted *rišt* ‘trigger’ from Omani Arabic *rīšah*; [Lonnet 2011](#)).

It is also worth noting that the Arabic ending *-(h)* is replaced by its MSAL equivalent when the noun is in the construct state, that is, final *-t* reappears. This would also happen in Arabic, but the alteration in the quality of the vowel is a clear signal that the suffix is to be considered an MSAL morpheme. Consider the following example from Morris’ Baṭḥari recordings:

- (1) Baṭḥari⁴
 mṣayš-it-həm bəss mṣayš-it-həm ḥawla fār ḥāmis
 sustenance-F.CS-3PL.M only sustenance-F.CS-3PL.M once only turtle
 w-ṣayd šālā mṣayš-ah ḥawil
 and-fish nothing sustenance-F once
 ‘Their sustenance, only that! Their sustenance once was only turtle and
 fish, there was nothing to eat in the past.’

The word *mṣayšah* ‘sustenance, food’ is a loanword from Arabic (as the *-ah* ending suggests). When suffixed with the possessive 3PL.M pronoun *-həm*, however, Baṭḥari *-it* replaces *-ah/-at* (note also, in the example, the use of the restrictive adverbial particle *bəss* ‘only’, which is a well-integrated loan from dialectal Arabic and occurs in alternation with Baṭḥari *fār*).

Finally, in Baṭḥari the Arabic definite article *(a)l-* is occasionally used instead of the MSAL definite article *a-*: *bə-l-ḥarifēt* ‘during the rainy season’.

⁴Audio file ‘20130929_B_B02andB04_storyofcatchingturtle’ recorded, transcribed and kindly shared with Fabio Gasparini by Miranda Morris. The recording was produced in the context of Morris’ and Watson’s “Documentation and Ethnolinguistic Analysis of the Modern South Arabian languages” project, funded by the Leverhulme Trust. More recordings are accessible at the ELAR archive of SOAS University of London. The transcription has been adapted.

3.2.2 Pronouns

The influence of Arabic can be observed, to an extent, even in the pronominal system, especially in those MSAL that are more exposed to contact due to the limited size of their speech communities. Lonnet (2011), for instance, reports how, despite the fact that in the MSAL the first person suffix pronoun is normally an invariable *-i*, in Ḥarsusi this can be replaced by *-ni* after verbs and prepositions (as is the case with Arabic; see also §3.3 for another interesting example concerning the marking of pronominal direct objects).

In addition, Baṭḥari relative pronouns (SG: *l-*, *lī* PL: *əlli*) are close to their equivalent in Janaybi Arabic (and diverge from the rest of MSAL, where a *ḍ*-element can be found). Baṭḥari has also borrowed the reflexive pronoun *ʕamr-* ‘oneself’ from the Arabic dialect of the Janaba, despite the existence of an original Baṭḥari term with the same meaning, *ḥanef-* (note that both terms must always be followed by a suffix personal pronoun). *ʕamr-* has also been given a plural form in Baṭḥari, based on MSAL derivational patterns, *ḥaʕmār-* (Morris 2017: 14).

3.2.3 Baṭḥari verbal plural marker *-uw*

Baṭḥari differs from the rest of the MSAL in that all 2/3PL.M verbal forms are marked by an *-uw* suffix, while in the other languages of the group these persons are marked by a *-Vm* ending and/or by internal vowel change (e.g. Mehri and Ḥarsusi *-kə(u)m* for the 2PL.M and *-ə(u)m*/umlaut for the 3PL.M of the perfective conjugation; *t-...-ə(u)m* and *y/i-...-ə(u)m* respectively for 2 and 3PL.M of the imperfective conjugation; Simeone-Senelle 2011: 1093–1094).

The origin of this suffix is uncertain. Its presence might well be connected to contact with Arabic (neighboring dialects have an *-u* or *-ūn* suffix in the 3PL.M person of the verb in both the perfective and imperfective conjugation) or to otherwise unattested stages of development internal to the MSAL verbal system. In this regard, Rubin (2017: 5) suggests for Mehreyyet the presence of a subjacent *-ə-* in 2nd/3rd plural masculine verbal suffixes which could therefore be somehow related to the Baṭḥari *-uw* marker. However, the optional simultaneous presence of apophony within the stem of 3PL.M verbal forms (similarly to what happens elsewhere in the MSAL), together with scarcity of data, prevents any conclusive assessment of the topic.

Simone Bettega & Fabio Gasparini

3.3 Syntax

At present, the syntax of the various MSAL has not been made the object of detailed investigation. The only scientific work dealing with this topic is Watson's (2012) in-depth analysis of Mehri syntax. However, Watson's thorough description provides only sporadic insights into the issue of language contact (as for instance the use in Mahriyōt, the Eastern Yemeni variety of Mehri, of a *swē ~ amma... yā* construction to express polycoordination, probably to be regarded as the result of Arabic influence; Watson 2012: 298). In general, though, the topic is left unaddressed in the literature, and more research is needed.

An interesting example of Arabic influence on MSAL syntax comes from Gasparini's data on Baṭḥari. In Baṭḥari, as in the other MSAL, pronominal direct and indirect objects may require a particle *t-* to be inserted between them and the verb, depending on the morphological form of the verb itself. Masculine singular imperatives, for instance, require the presence of the marker, as shown in the following example:

- (2) Baṭḥari (Gasparini, unpublished data):

zum t-ī t-ih
give.IMP ACC-1SG ACC-3SG.M
'Give it to me.'

Example (3), in contrast, shows that the pronominal indirect object *-(ə)nī* is suffixed directly to the verb as it would be in Arabic (see §3.2.2).

- (3) Baṭḥari (Gasparini 2018: 66):

zəm-ənī θrɛh
give.IMP-1SG two.M
'Give me both (of them).'

In other words, the introduction of the Arabic form of the object pronoun has caused the Baṭḥari object marker to disappear. Note that informants judged the alternative construction *zum t-ī θrɛh*, (with the use of the object marker *t-* and the 1SG object pronoun marker *-ī*) to be acceptable, but this form was not produced spontaneously.

A peculiarity of the MSAL spoken in Oman is the use of circumstantial qualifiers, a type of clausal subordination well attested in Gulf Arabic (Persson 2009). Baṭḥari regularly introduces predictive and factual conditional clauses *asyndetically* by using the structure [SBJ.PRO w-SBJ.PRO]. Consider (4):

- (4) Baṭḥari (Gasparini, unpublished data)

hēt w-hēt aṣbaḥ-k aḥayr saḥīr-e t-ōk
 2SG.M and-2SG.M wake_up.PRF-2SG.M better brand.PTCP-PL.M ACC-2SG.M
 lā w ham aṣbaḥ-k aḥass hāmā-k? w-maraḏ zēd
 NEG and if wake_up.PRF-2SG.M worse hear.PRF-2SG.M and-illness huge
 l-ōk nhā saḥīr-e t-ōk škīl-e t-ōk
 to-2SG.M 1PL brand.PTCP-PL.M ACC-2SG.M scar.PTCP-PL.M because_of
 mən a-gab
 DEF-infection.

‘In case you wake up feeling better / (we) do not brand you but in case you wake up feeling worse / do you understand? And you are seriously ill / we brand you and scar you because of the infection.’

The first clause *hēt w-hēt aṣbaḥ-k aḥayr* is an asyndetical circumstantial qualifier functioning as a predictive conditional clause. It contrasts with *w ham aṣbaḥ-k aḥass*, in which the conjunction *ham* introduces a counterfactual conditional clause.

In Omani Mehri conditional clauses are commonly introduced through conjunction of pronouns (Watson et al. forthcoming: 211). This structure is unattested in Yemeni Mehri:

- (5) Mehri (Watson et al. forthcoming: 211)

sēh wa-sēh t-ḥam-ah lā ḥib-sa yi-kaḷ-am
 3SG.F and-3SG.F 3SG.F-want.IMPF-3SG.M NEG parents-3SG.F 3M-let.IMPF-PL
 t-ēs ta-ghōm š-ih lā
 ACC-3SG.F 3SG.F-go.SBJV with-3SG.M NEG

‘If she doesn’t want him, her parents won’t let her go with him.’

These uses closely resemble those of Gulf Arabic, where circumstantial qualifiers are widely attested to codify predictive and factual conditional and consecutive clauses.

3.4 Lexicon

In the case of the MSAL, it can often be difficult to clearly set apart the effects of Arabisation from those of modernisation and lifestyle changes (which is not surprising, since the two phenomena are interrelated). According to what the speakers themselves report,

Simone Bettega & Fabio Gasparini

it was only since the introduction of formal education, and the awareness of [Modern Standard Arabic] via the media, that Arabic became the second language for many of the MSAL speakers in Dhofar, and, in the case of younger speakers, often to the detriment of their proficiency in their MSAL variety (Davey 2016: 11).

As a consequence, phenomena of borrowing (such as codeswitching and loan-words) are particularly common, especially in those varieties (and in the idiolects of individuals) which are more exposed to Arabic. The following is a good example of codeswitching in Baṭḥari (note that the speaker in question tended to employ Janaybi forms more than other informants):

- (6) Baṭḥari (Gasparini, unpublished data)
- | | | |
|----------------------|-------------------------------|------------------------|
| mēt | məssəlīm nə-šāhəd | l-ōk |
| die.PRF.3SG.M muslim | 1PL-say_šahada.IMPF for-2SG.M | |
| w-y-sabbah-uw | w-y-kabbər-uw | w-y-hālul-uw |
| and-3M-pray.IMPF-PL | and-3M-pray-PL | and-3M-praise_allah-PL |
- ‘(If) a muslim dies / we say the *šahada* for you / and they pray and say ‘*allāhu ʔakbar*’ and praise Allah.’

In (6) the speaker makes use of several Arabic verbs related to the semantic field of religious practices, which are not lexically encoded in Baṭḥari. This might indirectly show the introduction of new ritual practices at a certain point of the history of the tribe. Note that C₂-geminate stems such as *ysabbahuw* and *ykabbəruw* represent verbal patterns not attested in MSAL morphology, and are therefore easily identifiable as loans.

Morris (2017: 15) makes the important remark that lexical erosion is directly connected with the loss of importance of a language in the eyes of its speakers. She gives the example of the Baṭḥari word for ‘home, living quarters’, for which speakers nowadays frequently resort to some version of Arabic *bayt*, while the many possible original synonyms are falling into disuse. Many of these (*kədōt*, *mōhen*, *mašʕar*, *mōḥayf*, and *ḥader*) are connected to traditional ways of living which have all but disappeared in the course of the last 40–50 years, so that speakers probably judge them inadequate to refer to modern built houses.

3.4.1 Numerals

Watson (2012: 3) reports that “[w]hile Mehri cardinal numbers are typically used for both lower and higher cardinals in Mehreyyet, Mahriyōt speakers, in common

with speakers of Western Yemeni Mehri, almost invariably use Arabic numbers for cardinals above 10.” This type of lexical substitution connected to numerals higher than ten is also mentioned by [Lonnet \(2011\)](#) and [Simeone-Senelle \(2011: 1088\)](#), who states that “[n]owadays the MSAL number system above 10 is only known and used by elderly Bedouin speakers.” [Watson & Al-Mahri \(2017: 90\)](#) note that it is mostly younger generations (especially in urban settings) who have lost the ability to count beyond ten. Interestingly, they point out that telephone numbers are given exclusively in Arabic, “possibly due to the lack of a single-word MSAL equivalent to Arabic *ṣufr* ‘nothing’.”

3.4.2 Spatial reference terms

According to [Watson & Al-Mahri \(2017: 91\)](#) the MSAL employ topographically variable absolute spatial reference terms. In other words, these terms can differ depending on the language employed, the moment of the utterance and the position of the speaker in relation to absolute points of reference. For instance, in and around the city of Salalah in Dhofar, both in the mountains and on the coastal plain, the equivalents of the words for ‘sea’ and ‘desert’ are used to indicate south and north, respectively, in both Mehri (*rawram* and *nagd*) and Šehri (*ramnam* and *fagir*). This is because the sea lies to the south and the desert lies to the north (beyond the mountains). In other parts of the coastal plain, however, the word for ‘mountains’ (*šher*) is used to indicate north instead. Another common way to describe south and north is to refer to the direction in which the water flows, with the result that the same word that means ‘south’ on the sea-side of the mountains can be used to indicate ‘north’ on the desert-side. However, all these rather complex sets of terms are being rapidly replaced, particularly in the speech of the younger generations and among urban populations, with the Arabic words for south and north (*ḡanūb* and *šimāl* respectively).

3.4.3 Colour terms

The MSAL lexically encode four basic colour terms: white, black, red and green ([Bulakh 2017: 261–262](#)). For example, in Šehri one can find *lūn* for ‘white’, *hər* for ‘black’, *ʕəfər* for ‘red’ (and warm colours in general, including brown) and *šəzrər* for ‘green’ (and everything from green to blue). A fifth colour term, *šəfrər* ‘yellow’ (Mehri *šāfər*), is most probably an adapted borrowing from Arabic already present at the common MSAL level ([Bulakh 2017: 271](#)).

A preliminary field inquiry on the subject was conducted by Gasparini in 2017, with 6 young speakers from the city of Salalah and its immediate surroundings,

Simone Bettega & Fabio Gasparini

all between 20 and 35 years old and all bilingual in Ş̤ehri and Arabic. The results of the tests showed a remarkable degree of idiolectal variation in the colour labeling systems employed by the informants, with different levels of interference from Arabic. Remarkably, when asked to label colours in Ş̤ehri from a printed basic colour wheel, which was shown to them during interviews, all the speakers used the Arabic word for ‘blue’, *azraq*, which seems to have replaced *ś̤əḏr̥r̥* (traditionally used for both blue and green, but now confined to the latter). Two speakers also used *aḥḍar* for ‘green’, claiming that they could not recall the Ş̤ehri term. In addition, only one speaker used *ʕəf̥ar* for ‘brown’, Arabic *bunnī* being preferred by the other interviewees. The three basic colours ‘white’, ‘black’ and ‘red’, however, were regularly referred to using the Ş̤ehri forms by all speakers. Summing this up, it would seem that the Ş̤ehri colour system (at least in urban environments, but see below) is undergoing a radical process of restructuring. The three typologically fundamental colour terms are retained in most contexts, and a distinction between blue and green is being introduced through reduction of the original semantic spectrum of *ś̤əḏr̥r̥*, adoption of the Arabic word for blue, and subsequent replacement of *ś̤əḏr̥r̥* with *aḥḍar* (which indicates only green in Arabic). Further distinctions are either being replaced with the corresponding Arabic terms, or introduced if not part of the original semantic inventory of the language.

On this matter, Watson & Al-Mahri (2017: 90) argue that colour terms (together with numbers) are often among the first lexical items to be lost in contexts of linguistic endangerment, and that this is precisely the case with the MSAL. They write that even children in rural communities are now employing Arabic terms to refer to the different breeds of cattle (which traditionally used to be referred to by use of the three basic colour terms ‘white’, ‘black’ and ‘red’). This is probably a result of the fact that even in villages younger generations are no longer involved in cattle herding. Examples include *aḥmar* ‘bay’ in place of Mehri *ōfar* or Ş̤ehri *ʕofer*, *aswad* ‘black’ in place of Mehri *ḥōwar*, and *abyaḏ* ‘white’ in place of Mehri *ūbōn*.

3.4.4 Other word classes

Watson & Al-Mahri (2017: 90) note that, since the introduction of a public school system in Arabic in the 1970s, a number of common lexical items and expressions in Mehri and Ş̤ehri have been replaced by the corresponding Arabic ones. Lonnet (2011) also remarks that borrowings from Arabic are particularly common among particles and function words. Examples include *naḥs aš-šī* ‘the same thing’ for Ş̤ehri *gens*, Mehri *gans*; *lākin* ‘but’ in place of Ş̤ehri *duʰn* and *min*

du^hn, Mehri *lahinnah*; *yaʃnī* ‘that is to say’ and *ʃabārah* in place of Šəhri *yaḥīn*, Mehri (y)*aḥah*; *tamām* ‘fine’ in place of Šəhri *ḥayšōf* and Mehri *hīs taww* ~ *histaww*; Mehri and Ḥarsusi vocative *yā* ‘oh’ in place of MSAL *ʔā*–; Šəhri *bdan*, Mehri *ʔabdan* ‘never, not at all’, against Mehri and Ḥarsusi *bəhawʔ*, Šəhri *bhoʔ*. Consider also the case of Arabic *bəss* ‘only’, already mentioned in §3.2.1. In Mehri as in Baṭṭari, this particle appears now to be interchangeable with its equivalent *ār*, as example (7) shows:

- (7) Mehri (Sima 2009: 328, cited in Watson 2012: 371; transcription adapted)
 bass ta-ʔʃam-h kād aḥah ār ʔʃām
 only 2SG.M-taste.IMPF-3SG.M INT fine only 2SG.M-taste.IMPF
 ǝ-maḥḥ
 of-clarified_butter
 ‘Just taste it, like it is just the taste of clarified butter.’

As is predictable, also in this field Baṭṭari is the language most affected by Arabic: besides those already cited, we might add the expressions *zēn* ‘well’, (*a*)*barr* ‘outside’ (also in Mehri, as opposed to Soqotri *ter*), *ḥalaṣ* ‘and this is it’ (used to end a narrative). Finally, Watson (2012: 3) remarks how “Mahriyōt also exhibits structures unattested in Mehreyyet such as ‘What X!’ phrases reminiscent of Arabic, e.g. *maṭwalk* ‘How tall you m.s. are!’”.

4 Conclusions

Throughout this chapter we have repeatedly pointed out how research on the MSAL, and in particular on the effects that contact with Arabic has had on their evolution, is still far from reaching its mature stage. Much remains to be done, in particular, in terms of sheer documentation, especially in the case of the most endangered varieties (Hobyōt, Ḥarsūsi, Baṭṭari). In addition to this, and although Watson’s (2012) work has greatly contributed to expanding our knowledge in this area, MSAL syntax remains a strongly neglected field of inquiry. Finally, our knowledge of the history of the MSAL prior to the twentieth century (and therefore the history of their contact with Arabic) is extremely poor.

It must also be remarked that, although the most widely spoken among the MSAL are undoubtedly better documented, very little is known about the effects that urbanisation has had on their speech communities in recent years. In particular, anecdotal evidence suggests that the varieties of Šəhri and Soqotri spoken in Salalah and Ḥadibo are undergoing rapid change under the influence

Simone Bettega & Fabio Gasparini

of Arabic (both the standard variety of the language, which children learn in school, and the dialects). Fieldwork conducted in the two abovementioned urban centres could provide extremely valuable information concerning the effects of contact between Arabic and Modern South Arabian.

Despite the far-from-complete state of the research in this field, what we currently know is sufficient to say that contact has had a strong impact on the MSAL. Though this is more evident in the area of lexicon, where borrowings are legion, phonetics and phonology have also been affected (though to a different extent from one language to another). Morphology and syntax, on the contrary, appear to be more resistant to contact-induced change, though in the most endangered varieties one can notice a partial disruption of the original pronominal system and verbal paradigm, and though the seemingly high degree of resistance to external influence shown by MSAL syntax could actually be due to our limited knowledge of the subject.

One last note is due concerning another heavily neglected topic, namely the effects that contact with the MSAL have had on spoken Arabic. Though we have not addressed the question in the course of this paper, evidence drawn from the existing literature [Simeone-Senelle \(2002\)](#) suggests that this influence, too, is not completely absent, and that further research in this direction could produce interesting results.

Further reading

[Morris \(2017\)](#) can be thought of as a general introduction to contact between MSAL and Arabic.

[Watson & Al-Mahri \(2017\)](#) offer an intriguing account of how language change, contact with Arabic and modification of the traditional environment are all deeply interrelated.

[Lonnet \(2011\)](#) – although limited in scope and extension due to its nature as an encyclopedic entry – offers interesting highlights on the effects of contact on the MSAL.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person		
ACC	accusative	MSAL	Modern South Arabian
CS	construct state	PTCP	participle
F	feminine	PRO	pronoun
IMP	imperative	PRF	perfect
IMPF	imperfect	PL	plural
INT	intensifier	SG	singular
M	masculine	VOC	vocative
NEG	negative		

References

Beeston, Alfred F. L. 1981. Languages of pre-Islamic arabia. *Arabica* 28(2/3). 178–186.

Bulakh, Maria. 2017. Color terms of the modern south arabian languages: a diachronic approach. In Leonid Kogan, Natalia Koslova, Sergey Loesov & Sergey Tishchenko (eds.), *Babel und bibel 1: annual of Ancient Near Eastern, Old Testament and Semitic studies*, 269–282. Winona Lake, IN: Eisenbrauns.

Davey, Richard J. 2016. *Coastal Dhofari Arabic: A sketch grammar*. Leiden: Brill.

Dufour, Julien. 2016. *Recherches sur le verbe sudarabique moderne*. Paris: Ecole pratique des hautes études habilitation dissertation.

Edzard, Lutz. 2017. On the role of modern South arabian within a comparative Semitic lexicographical project. *Brill’s Journal of Afroasiatic Languages and Linguistics* 9(1–2). 119–138.

Gasparini, Fabio. 2017. Phonetics of emphatics in bathari. In Simone Bettega & Fabio Gasparini (eds.), *Linguistic studies in the Arabian Gulf*, 69–85. Turin: Dipartimento di Lingue e Letterature Straniere e Culture Moderne, Università di Torino.

Gasparini, Fabio. 2018. *The Baṭḥari language of Oman: Towards a descriptive grammar*. Naples: Università degli studi di Napoli “L’Orientale” doctoral dissertation.

Hoyland, Robert G. 2001. *Arabia and the Arabs: From the Bronze Age to the coming of Islam*. London: Routledge.

Kogan, Leonid. 2015. *Genealogical classification of Semitic: The lexical isoglosses*. Berlin: Walter de Gruyter.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world’s languages*. Oxford: Blackwell.

Simone Bettega & Fabio Gasparini

- Lonnet, Antoine. 2011. Modern South Arabian. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Morris, Miranda J. 2017. Some thoughts on studying the endangered modern South arabian languages. *Brill's Journal of Afroasiatic Languages and Linguistics* 9(1–2). 9–32.
- Naumkin, Vitaly & Victor Porkhomovsky. 1981. *Očerki po etnolingvistike sokotry*. Moscow: Nauka.
- Persson, Maria. 2009. Circumstantial qualifiers in gulf Arabic dialects. In Bo Isaksson, Heléne Kammensjö & Maria Persson (eds.), *Circumstantial qualifiers in Semitic: The case of Arabic and Hebrew*, 206–286. Wiesbaden: Harrassowitz.
- Porkhomovsky, Victor. 1997. Modern South arabian languages from a Semitic and hamito-Semitic perspective. In Janet Starkey (ed.), *Proceedings of the seminar for Arabian Studies vol. 27*, 219–223. Oxford: Archaeopress.
- Retsö, Jan. 2000. *Kaškaša, t-passives and the ancient dialects in arabia. Oriente Moderno* 19(1). 111–118.
- Ridouane, Rachid & Cédric Gendrot. 2017. On ejective fricatives in Omani mehri. *Brill's Journal of Afroasiatic Languages and Linguistics* 9(1–2). 139–159.
- Robin, Christian-Julien. 1991. La pénétration des arabes nomades au yémen. *Revue des mondes musulmans et de la Méditerranée* 61(1). 71–88.
- Rubin, Aaron D. 2014. *The Jibbali (Shaḥri) language of Oman: Grammar and texts*. Leiden: Brill.
- Rubin, Aaron D. 2017. A re-analysis of mehri grammar. Paper presented at the *Cinquièmes journées d'études sur les langues sudarabiques modernes*, September 2017, Paris.
- Sima, Alexander. 2009. *Mehri-Texte aus der jemenitischen Sharqīyah: Transkribiert unter Mitwirkung von Askari Hugayran Saad*. Wiesbaden: Harrassowitz.
- Simeone-Senelle, Marie-Claude. 2002. L'arabe parlé dans le mahra (yémen). In Otto Jastrow, Werner Arnold & Hartmut Bobzin (eds.), *"Sprich doch mit deinen Knechten aramäisch, wir verstehen es!": 60 Beiträge zur Semitistik: Festschrift für Otto Jastrow zum 60. Geburtstag*, 669–684. Wiesbaden: Harrassowitz.
- Simeone-Senelle, Marie-Claude. 2011. Modern South arabian. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1073–1113. Berlin: De Gruyter Mouton.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

12 *Modern South Arabian Languages*

- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Versteegh, Kees. 2014. *The Arabic language*. Second edition. Edinburgh: Edinburgh University Press.
- Watson, Janet C. E. 2011. South arabian and yemeni dialects. *Salford Working Papers in Linguistics and Applied Linguistics* 1. 27–40.
- Watson, Janet C. E. 2012. *The structure of Mehri*. Wiesbaden: Harrassowitz.
- Watson, Janet C. E. 2018. South arabian and arabic dialects. In Clive Holes (ed.), *Arabic historical dialectology: linguistic and sociolinguistic approaches*, 316–334. Oxford: Oxford University Press.
- Watson, Janet C. E. & Abdullah Musallam Al-Mahri. 2017. Language and nature in dhofar. In Simone Bettega & Fabio Gasparini (eds.), *Linguistic studies in the Arabian Gulf*, 87–103. Turin: Dipartimento di Lingue e Letterature Straniere e Culture Moderne, Università di Torino.
- Watson, Janet C. E., Abdullah Musallam Al-Mahri & Ali Al-Mahri. Forthcoming. *Taghamk afyat: A course in Mehri*. Wiesbaden: Harrassowitz.
- Watson, Janet C. E. & Alex Bellem. 2010. A detective story: emphatics in mehri. In Janet Starkey (ed.), *Proceedings of the seminar for Arabian Studies vol. 40*, 345–355. Oxford: Archaeopress.
- Watson, Janet C. E. & Alex Bellem. 2011. Glottalisation and neutralisation in yemeni Arabic and mehri: an acoustic study. In Majeed Hassan Zeki & Barry Heselwood (eds.), *Instrumental studies in arabic phonetics*, 235–256. Amsterdam: John Benjamins.
- Watson, Janet C. E. & Barry Heselwood. 2016. Phonation and glottal states in modern South arabian and San’ani Arabic. In Youssef A. Haddad & Eric Potsdam (eds.), *Perspectives on arabic linguistics XXVIII: papers from the annual symposium on arabic linguistics, gainesville, florida, 2014*, 3–36. Amsterdam: John Benjamins.
- Watson, Janet C. E., Bonnie Glover Stalls, Khlaïd Al-Razihi & Shelagh Weir. 2006. The language of jabal rāziḥ: arabic or something else? In Rob Carter & St John Simpson (eds.), *Proceedings of the seminar for Arabian Studies vol. 36*, 35–41. Oxford: Archaeopress.
- Wellsted, James R. 1837. Narrative of a journey into the interior of omán, in 1835. *Journal of the Royal Geographical Society of London* 7. 102–113.
- Winford, Donald. 2005. Contact-induced change: classification and processes. *Diachronica* 22(2). 373–427.
- Zammit, Martin R. 2011. South arabian loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Chapter 13

Northern Domari

Bruno Herin

Inalco, IUF

This chapter provides an overview of the linguistic outcomes of contact between Arabic and Northern Domari. Northern Domari is a group of dialects spoken in Syria, Lebanon, Jordan and Turkey. It remained until very recently largely unexplored. This article presents unpublished first-hand linguistic data collected in Lebanon, Syria, Jordan and Turkey. It focuses on the Beirut/Damascus variety, with references to the dialects spoken in northern Syria and southern Turkey.

1 Current state and historical development

Domari is an Indic language spoken by the Doms in various countries of the Middle East. The Doms are historically itinerant communities who specialize in service economies. This occupational profile led the lay public to call them the Middle Eastern Gypsies. Common occupations are informal dentistry, metalwork, instrument crafting, entertainment and begging. Most claim Sunni Islam as their religion, with various degrees of syncretic practices. Although most have given up their semi-nomadic lifestyle and settled in the periphery of urban centers, mobility is still a salient element in the daily lives of many Doms.

The ethnonym Dom is mostly unknown to non-Doms, who refer to them with various appellations such as *nawar*, *qurbāt* or *qarač*. The Standard Arabic word *yağar* for ‘Gypsy’ is variably accepted by the Doms, who mostly understand with this term European Gypsies. All these appellations are exonyms and the only endonym found across all communities is *dōm*. Only the Gypsies of Egypt, it seems, use a reflex of *yağar* to refer to themselves.

From the nineteenth century onwards, European travellers reported the existence of Domari in the shape of word lists collected in the Caucasus, Iran, Iraq



Bruno Herin

and the Levant (see [Herin 2012](#) for a discussion of these sources). The first full-length grammatical description of a dialect of Domari is by [Macalister \(1914\)](#), who described the dialect spoken in Palestine in the first years of the twentieth century. At present, the language is known to be spoken in Palestine, Jordan, Lebanon, Syria and Turkey. No recent account can confirm that it is still spoken in Iraq and Iran. There are roughly two dialectal areas: Southern Domari, spoken in Palestine and Jordan, and Northern Domari, spoken in Lebanon, Syria and southern Turkey. This geographical division is not clear cut because I have recorded speakers of Southern varieties in Lebanon and speakers of Northern dialects in Jordan. The main isogloss separating these two groups is the maintenance of a two-way gender system. Southern dialects have maintained the gender distinction, whereas it has mostly disappeared in the north. Compare Northern *gara* ‘(s)he went’ vs. Southern *gara* ‘he went’ and *garī* ‘she went’. These are sufficiently different to allow us to posit an early split. Mutual intelligibility appears to be very limited. A case in point is kinship terminology, which is largely divergent in both groups. Within Northern Domari, the Beirut/Damascus dialect stands out because of the glottal realization [ʔ] of etymological /q/ and the loss of the differential subject marker *-ən*.

No general statement can be made about language endangerment. Jerusalem Domari is reported to have only one fluent speaker left (Matras, this volume), but the presence of speakers of Palestinian Domari in other places may not be excluded. Young fluent speakers of Southern dialects are easy to find in Jordan. As far as Northern Domari is concerned, the language is no longer transmitted to the young generation in Beirut but it is in Damascus. In northern Syria, intergenerational transmission is quite solid. The situation in southern Turkey is, according to some consultants, more precarious, but I have personally witnessed quite a few children fully conversant with the language. In any case, bilingual Doms acquire both languages in early childhood, making both languages equally “dominant” in Van Coestem’s (1988; 2000) terms.

Many Dom groups are also found in Eastern Anatolia. These groups have shifted to Kurdish but maintained an in-group lexicon based on Domari, locally called Domani. According to what I could personally observe on the ground and what well-informed local actors reported to me, full-fledged Domari is not spoken beyond Urfa. East of Urfa, the shift to Kurdish is complete and even the in-group lexicon is only remembered by elderly individuals.

There are no reliable figures on the number of speakers of Domari. The language has often been mistaken for a variety of Romani but this claim has no linguistic grounds, except that they are both classified as Central Indo-Aryan Languages

with a possible Dardic adstrate.

2 Contact languages

Besides a Central Indic core and a Dardic adstrate, the language exhibits various layers of influence. Easily identifiable sources of contact are Persian, Kurdish, Turkish and finally Arabic. This suggests, quite logically, that the ancestors of the Doms left the Indian subcontinent, and then travelled into Persian-speaking lands, before reaching Kurdish- and Turkish-speaking areas (most probably in eastern Anatolia), before venturing into Arab lands. It is striking to see that the Iranian and Turkic elements in Domari are not uniform across Northern and Southern varieties, which suggests an early split in eastern Anatolia between speakers of both groups. The impact of Arabic is also not uniform both across Southern and Northern Domari, nor even within Northern Domari. What this means is that the validity of any discussion of the Arabic component in Domari is limited to the varieties considered.

The Beirut/Damascus dialect is undoubtedly the most arabized one within the Northern group, pointing to an earlier settlement of the community in an Arabic-speaking environment. Bilingualism (Domari–Arabic) is general in Lebanon and Syria. Except perhaps for very young children who have not yet acquired any other language, monolinguals in Domari are not to be found.

As far as Turkey is concerned, trilingualism in Domari, Turkish and Kurdish is not uncommon, especially in southern Turkey around Gaziantep. In the Hatay province, many speakers above the age of forty are trilingual Domari–Arabic–Turkish. The generations born here in the eighties onwards did not acquire Arabic.

According to personal recollection from various consultants, the Beirut/Damascus community used to spend the winter in Lebanon, and would go back to Damascus in the summer. This semi-nomadic way of life seems to have stopped when the civil war in Lebanon began. Although movements between Beirut and Damascus remained frequent, this phenomenon ceased to be seasonal. In Damascus, they settled in the area of Sayyida Zaynab, in the suburbs of the city, and in Beirut many of them settled in Sabra. Since the civil war started in Syria, virtually all the Damascus community have moved to Lebanon and settled in refugee camps in the Bekaa Valley close to the Syrian border.

Bruno Herin

3 Contact-induced changes in Northern Domari

As noted above, Domari speakers in Lebanon and Syria are also fully proficient in Arabic, to the point that I have never encountered or heard of any monolingual adult. The Dom community, although largely endogamous and socially isolated, cannot afford monolingualism, primarily because of their peripatetic profile. As far as one can judge, their proficiency in Arabic is that of any monolingual native speaker of Arabic. Their pronunciation, however, is often not fully congruent with the local dialect spoken in the immediate vicinity of their settlements. This is, as usual, due to the variety of inputs and migration after acquisition. The Doms of Beirut for instance, do not speak Beirut Arabic and their speech is immediately perceived as Syrian by Lebanese because they do not raise /ā/. Raising of /ā/ towards [e:] is the hallmark of Lebanese Arabic in perceptual dialectology. Proficient speakers of Domari all exhibit Arabic–Domari bilingualism. On the whole, there is a general license to integrate any Arabic lexeme in Domari speech, even when a non-Arabic morpheme exists. Code-switching is also very common and there seems to be no conservative ideology about linguistic practices, leading to a very permissive environment for language mixing.

3.1 Phonology

All the segmental phonology of Arabic has made its way into Domari. Arabic stands out cross-linguistically because of its series of back consonants such as the pharyngeals /ħ/ and /ʕ/, the post-velars /q/, /ħ/ and /ɣ/, and a set of velarized consonants whose number varies from dialect to dialect. Typically, sedentary varieties in the Levant minimally exhibit contrast between /d̥/, /z̥/, /t̥/ and /s̥/. In Domari, the pharyngeals /ħ/ and /ʕ/ are commonly found in loans from Arabic: *ħdər* *h-* ‘watch’ (from Levantine Arabic *ħiḍir* ‘he watched’). The same goes for /ʕ/: *ʕammər* *kar-* ‘build’ (from Arabic *ʕammār* ‘he built’). An oddity surfaces in the word for coffee, realized *ʔaħwa* from Arabic *ʔahwe*. These pharyngeals are also common in Kurdish-derived items such as *ħazār* ‘thousand’, *moʕōri* ‘ant’ and also in the inherited (Indic) stock in *ʕaqqōr* ‘nut’. Post-velar /q/, /ħ/ and /ɣ/ are found in all the layers of the language: *qāla* ‘black’ (inherited), *qāpī* ‘door’ (Turkish), *sāy* ‘alive’ (Kurdish), *yarīb* ‘strange’ (Arabic). The most striking innovation of the Beirut/Damascus dialect is the glottal realization [ʔ] of /q/: *ʔər* ‘son’ (< *qər*), *ʔāyīš* ‘food’ (< *qāyīš*). This innovation is very likely contact-induced because it is commonly found in the Arabic dialects of both Damascus and Beirut and beyond.

Velarized consonants mostly surface in the Arabic-derived stock as in *naḍḍaḥ* *kar-* ‘clean’ (< Arabic *naḍḍaḥ* ‘he cleaned’), but also in pre-Arabic items: *dāwaṭ*

‘wedding’ (borrowed from Kurdish but ultimately from Arabic *daḡwa* ‘invitation’), *pāṣ* ‘at him’ (< Old Indo-Aryan *pārśvá* ‘side’). It is still unclear to what extent velarization in Domari continues Indo-Aryan retroflexion (Matras 2012: 64). Domari also kept a contrast between /p/ and /b/, not found in Arabic: *bīrōm* ‘I feared’ vs. *pīrōm* ‘I drank’.

As far as vowels are concerned, Levantine Arabic exhibits either a two-way distinction in the short vowel system (/a/ and /ə/) or a three-way distinction (/a/, /i/ and /u/). In Northern Domari, only the two short vowels /a/ and /ə/ are contrastive: *kārī* ‘house’ vs. *karī* ‘pot’. Such a paucity of contrastive short vowels is probably due to contact with Arabic varieties which exhibit a two-way system (/a/ vs. /ə/), such as many Lebanese and Syrian dialects. Most Arabic dialects in the area have a five-way system of long vowels because of the monophthongization of /ay/ and /aw/: /ā/, /ī/, /ū/, /ē/ and /ō/. In addition to these long vowels, Domari displays another contrast between /ā/ and a back /ā/ (IPA [ɑ:]): *māsī* [ma:si:] ‘meat’ (< Old Indo-Aryan *māmsá*) vs. *mās-i* [mɑ:si:] ‘month.PL’ (< Old Indo-Aryan *māsa*).

Domari has also preserved distinct suprasegmental features, such as final syllable stress assignment. Arabic-derived items are fully integrated into this pattern and bear final primary stress, whether common nouns or proper nouns: Domari [faˈdja] vs. Arabic [faˈdja] (personal name *Fādyā*). An interesting phenomenon is that Arabic epenthetic vowels in final-syllable position are reinterpreted as plain vowels and bear primary stress. Compare Domari [sˈaʔab] and Arabic [sˈaʔab]) ‘difficult’; Domari [waˈdʔaʔ] and Arabic [ˈwadʔəʔ] ‘situation’.

3.2 Morphology

Northern Domari has not borrowed any derivational or inflectional morphemes from Arabic. This is of course due to the fact that Arabic morphology is mostly non-concatenative. Borrowed morphology mostly comes from Kurdish and Turkish, whose morpheme segmentation is much more transparent. These morphemes must have entered Domari when Kurdish and Turkish were contact languages of Domari. A case in point is the Kurdish diminutive *-ək*, which has made its way into all layers of the lexicon: *panč-ək* ‘tail’, *ḡar-ək* ‘bone’ (both Indic), *taht-ək* ‘wood’, *qannīn-ək* ‘bottle’ (both derived from Arabic: *taht* ‘bed’ and *qannīne* ‘bottle’). The dialects of northern Syria and southern Turkey have also borrowed from Kurdish the comparative suffix *-tar*, the Turkish conditional marker *-sa* and the Turkish superlative marker *ān*. These constructions are not available in the dialect of Beirut/Damascus, which relies entirely on Arabic-derived material. Compare the translation of the Arabic sentence *inte aḡsan minni* ‘you are better

Bruno Herin

than me’ into Sarāqib Domari (1) and Beirut/Damascus Domari (2):

- (1) Sarāqib Domari
tō dēšōm bḥēz-tar ištōre
2SG 1SG.ABL good-CMPR COP.2SG
‘You are better than me.’
- (2) Beirut/Damascus Domari
tō aḥsan wēšōm ištōr
2SG better 1SG.ABL COP.2SG
‘You are better than me.’

Sarāqib is located in northern Syria and the dialect spoken by the Doms of Sarāqib is a good representative of the Domari of northern Syria and southern Turkey. Three differences are immediately apparent. The first is morphological, whereby there are different forms for the ablative of the first-person pronoun. The second difference is syntactic: in (1) the standard of comparison precedes the comparative adjective (*dēšōm bḥēz-tar*) and in (2) it follows it (*aḥsan wēšōm*).¹ The Beirut/Damascus Domari syntax exhibits full congruence with the Arabic syntax. The third difference is lexical. Because Beirut/Damascus Domari does not have at its disposal the morpheme *-tar*, speakers are obliged to draw on Arabic for the comparative. This phenomenon, labelled “bilingual suppletion” by Matras, is described at length for Jerusalem Domari (Matras 2012: 379–382; see also Matras, this volume: §3.5).

Beirut/Damascus Domari also relies entirely on Arabic material for the expression of time and date, as shown in (3). In northern Syria, speakers favour the use of inherited numerals, as exemplified in (4).

- (3) Beirut/Damascus Domari
mānane mi-s-sāʿa ʿašra la s-sāʿa sabʿa tmāne ōtanta sa
stay.IPFV.1PL from-DET-hour ten to DEF-hour seven eight there all
čāy-an-sa
children-OBL.PL-COM
‘We stay there with the all the kids from ten o’clock to seven or eight o’clock.’

- (4) Sarāqib Domari

¹Comparative constructions typically involve two noun phrases (NPs). Stassen (2013) labels the object of comparison the “compare NP” and the other the “standard NP”.

13 Northern Domari

ḥatta saʕat štār ēwar mānde ē čōrt-ə-ma
 until hour four evening stay.PFV.3SG DEM.OBL wasteland-OBL-IN
 ‘He stayed until 4pm in this wasteland.’

Some speakers of Beirut/Damascus Domari also extend the use of Arabic to higher numerals because, according to their own judgment, they have difficulties retrieving the pre-Arabic options. A look at their distribution reveals that the main parameter that triggers the use of Arabic items is not so much high numerals, but rather the complexity of the numeral. Compare in this regard (5) and (6). In (5), the speaker uses Arabic for the more complex numeral ‘95000’ but uses Domari items for simpler ‘2000’, ‘3000’ and ‘4000’.

- (5) Beirut/Damascus Domari
 pārda abōs šaʔʔ-āka ši ḥamse u tisīn alf dolar
 buy.PFV.3SG 3SG.BEN flat-INDF about five and ninety thousand dollar
 ‘He bought a flat for her, about ninety-five thousand dollars.’

- (6) Beirut/Damascus Domari
 načīš-a-ki dī ḥazār trən ḥazār štār ḥazār dfaʕ
 dancing-OBL-ABL two thousand three thousand four thousand pay
 kaštand dādōs kē
 do.PROG.3PL her.mother BEN
 ‘They give two, three, four thousand (dollars) to her mother from dancing.’

As noted above, it appears that the use of Arabic numerals is closely linked to language dominance. Speakers themselves are aware of it and when asked why they do not use Domari numerals, they justify it claiming a lack of proficiency. Looking at the distribution of inherited and Arabic numerals is therefore a good way to assess whether language attrition is incipient or not.

The impact of Arabic is also apparent in some morphological differences between the Beirut/Damascus variety and the dialects of northern Syria. For instance, the verb *sək-* means ‘learn’. The Beirut/Damascus dialect adds the passive suffix *-yā/-ī*. The corresponding verb in Arabic *tʕallam* is marked with the valency-decreasing prefix *t-*. What the speakers of Beirut/Damascus Domari have done is to replicate the valency-decreasing prefix *t-* of *tʕallam* by means of the Domari passive suffix *yā/-ī*: *skə-rd-ōm* (learn-PFV-1SG; northern Syria) vs. *sk-ī-r-ōm* (learn-PASS-PFV-1SG; Beirut/Damascus) ‘I learnt’.

Unlike Southern Domari, Northern Domari does not normally transfer Arabic plurals. Speakers simply use the singular form and add the Domari plural marker

Bruno Herin

-ī(n): *azʕar-īn* ‘thugs’ instead of the Arabic plural *zuʕrān*. Arabic plurals do surface at times, but only when they exhibit a high degree of independence within the lexicon. Examples are *ʔarāyb-ē-mā* (relatives-PL-1PL) ‘our relatives’, *ǧīrān-ē-mā* (neighbors-PL-1PL) ‘our neighbors’, from Arabic *qarāyib* and *ǧīrān*. Although these items have singular forms (respectively *qarīb* and *ǧār*), they are arguably lexicalized plurals and independent entries in the Arabic lexicon.

3.3 Syntax

3.3.1 Constituent order

The impact of Arabic in the realm of syntax is not uniform across Domari dialects. Dialects of northern Syria and southern Turkey show a strong tendency towards a head-final constituent-order typology, both within the NP and the clause. This feature is areal, so its presence in Domari may well be contact-induced. The canonical syntax of the NP is (demonstrative) (numeral) (adjective) (noun) noun. Complex NPs could only be retrieved through elicitation (examples (7) to (10)) and hardly occur in spontaneous speech. Example (7) illustrates the canonical syntax, where all the modifiers appear to the left of the head. Speakers of Beirut/Damascus Domari, however, tend to dislocate some modifiers to the right of the head, converging towards the Arabic syntax, as in (8), (9) and (10).

- (7) Sarāqib Domari
 ē štār lāfty-ən-ki dād-ō-sā
 DEM four girl-OBL.PL-ABL mother-SG-3PL
 ‘the mother of these four girls’
- (8) Beirut/Damascus Domari
 dād-ō-sā štār lāfty-an-ki
 mother-SG-3PL four girl-OBL.PL-ABL
 ‘the mother of the four girls’
- (9) Beirut/Damascus Domari
 nām-ē-sā ǧəwr-an-ki tərən-an-ki
 name-PL-3PL woman-OBL.PL-ABL three-OBL.PL-OBL
 ‘the names of the three girls’
- (10) Beirut/Damascus Domari
 dōm-an-sa ēr-an-sa štār-an-sa
 dom-OBL.PL-COM DEM-OBL.PL-COM four-OBL.PL-COM
 ‘with these four Doms’

In (9), the speaker also dislocates to the right the numeral *trən* ‘three’ which normally appears to the left giving the expected order *trən ġəwr-an-ki nām-ē-sā* (three woman-OBL.PL-ABL name-PL-3PL). The numeral remains unmarked for case when it appears to the left of the head. When it is placed to the right, it agrees in case with the head. This is also the case with the demonstrative in (10). Here the normal order would be *ē štār dōm-an-sa* (DEM four Dom-OBL.PL-COM). The fact that speakers replicate case marking on right-dislocated modifiers suggests that they feel the need to strengthen constituency in case of non-canonical syntax.

The influence of Arabic also surfaces in the Beirut/Damascus dialect in the syntax of the quantifier *sa* ‘all’. This is normally located to the right of the head: *ammāt sa* ‘all the people’ (‘people all’). In Beirut/Damascus, *sa* consistently surfaces to the left, like the Arabic quantifier *kull*: *sa ammat* (Arabic *kull in-nās*).²

3.3.2 Internal object

Domari speakers regularly replicate Arabic constructions and idioms, but tends to do so by recruiting inherited or pre-Arabic material – they do not borrow Arabic material. For instance, all dialects have replicated the so-called internal object construction, commonly used in Arabic as a predicate-modifying construction. Consider for instance (11) in Jordanian Arabic, where the speaker narrows the scope of the predication using the verbal noun *ʕirəf* ‘knowledge’, derived from the verb *ʕirif* ‘he knew’, and modifies it with the adjective *ṭayyib* ‘good’. In (12), the speaker has used the deverbal derivation *kūš* from the root *kū-* ‘throw’ and coded it as an object, as evident from the accusative marker *-əs*. This replicates the Arabic internal object construction as illustrated in (11).

(11) Jordanian Arabic

baʕrif-hum ʕirəf ṭayyib
know.IMPF.1SG-3PL knowledge good
‘I know them well.’

(12) Sarāqib Domari

dād-ōs ibnḥarām e ē kūš-əs
mother-3SG son.of.illicit COP DEM throwing-ACC
ktōs-s-e
throw.PFV.3SG-OBJ.3SG-PRS

‘His mother is heartless for having thrown (her baby) in such a way.’

²Arabic *kull* can also appear to the right as in *in-nās kull-ha ~ kull-hum* ‘all the people’ but this is a marked syntax.

Bruno Herin

3.3.3 Impersonal construction

Speakers also replicate the Arabic impersonal construction with the indefinite pronoun *il-wāḥad* by way of the inherited noun *mānās* ‘individual, people’. Example (13) illustrates the use of *il-wāḥad* in (Jordanian) Arabic. In (14), the sequence *gzare māns-as* corresponds to Arabic *biṣiḍḍ il-wāḥad*, literally ‘it bites one’. The fact that *māns-as* replicates *il-wāḥad* is also apparent from the accusative marking in Domari, which normally surfaces only with definite objects. The referent here is by nature indefinite and non-referential, so accusative marking in Domari can only be explained by the presence of the definite article *il-* in Arabic *il-wāḥad*.

- (13) (Jordanian) Arabic
 kān ʕēb il-wāḥad yrūḥ ʕala ʔutēl
 be.PRF.3SG.M shameful DEF-one go.IMPV.SBJV.2SG.M to hotel
 ‘One was ashamed to spend the night in a hotel.’

- (14) Beirut Domari
 aṣti ši hana lli baḥr-a-ma e gzare māns-as
 EXS too DEM REL sea-OBL-IN COP bite.IPFV.3SG man-ACC
 ‘There is this thing in the sea, it bites you.’

3.3.4 Auxiliaries

Probably the most striking difference between Southern and Northern Domari as far as the Arabic component is concerned is the absence of Arabic inflected material in the latter. Only the dialect of Beirut/Damascus has borrowed the auxiliaries *kān* (with its imperfect form *bikūn*), *ṣār* and *ḥalli*.

- (15) Beirut Domari
 ṣār ḡahhəz lakand lāfty-a kē bḥēr
 become.PRF.3SG prepare do.SBJV.3PL girl-OBL BEN well
 ‘They prepare the girl well now (for the wedding).’

- (16) Beirut Domari
 ḥaḍra kān məḡnār-a
 Khadra be.PRF.3SG.M breastfeed.IPFV.3SG-PST
 ‘Khadra was breastfeeding.’

- (17) Beirut Domari

13 Northern Domari

āwande bikūn krēnde mā kē kyāmōr
 come.IPFV.3PL be.IMPF.3SG do.PRF.3PL 1SG BEN something

‘(My kids) would come and they would have done something (naughty).’

In (15), the subject is in the 3PL but *šār* remains invariable, as the 3PL is *šāru*. In (16), the subject is feminine so if there was agreement one would expect *kānat*, not masculine *kān*. A further intriguing feature in (16) is the redundancy in past marking, first with *kān* and second with the past suffix *-a*, which in northern Syria and southern Turkey Domari suffices to mark past tense. The same invariability is apparent in (17) where the 3PL of *bikūn* should be *bikūnu*. These auxiliaries have the same semantic load as in Arabic. The morpheme *šār* puts emphasis on the inception of the event, *kān* followed by the imperfect places the event in the past and gives it an iterative/habitual aspect and *bikūn* describes a possible state of affairs not attested at the time of utterance. Arabic *šār*, *kān* and *bikūn* are absent in the dialects of northern Syria and southern Turkey. The only auxiliary that has been replicated here is *šār*. These dialects, however, have only replicated the structure, not the substance, that is they rely on inherited morphemes, as exemplified in (18). The speaker simply translates Arabic *šār* with the Domari equivalent *hər*, replicating the Arabic structure *šār* + subjunctive (see Manfredi, this volume). A further difference is word order, with the verb placed clause-finally in the subordinate clause.

- (18) Sarāqib Domari
 hər wārsinda lwār
 become.PFV.3SG rain hit.SBJV.3SG
 ‘It started raining.’

As noted above, in these dialects the functions of Arabic *kān* are expressed by the inherited past suffix *-a*. The functions covered by Arabic, *bikūn*, however do not seem to be encoded in the grammar of these dialects.

In Levantine Arabic, the imperative form *ħalli* ‘let’ of *ħalla* ‘he let’ is often used to soften an order and allows the speaker to avoid using an imperative, flagging a suggestion or an invitation, as shown in (19):

- (19) Jordanian Arabic
 ħalli ibn-ak yrūħ la ġ-ġēš
 let son-2SG.M go.IMPF.3SG.M to DEF-army
 ‘Let your son serve in the army.’

Bruno Herin

This auxiliary has been borrowed into Beirut/Damascus Domari with the exact same function, as illustrated in (20). In this case too, *ħalli* remains invariable and does not surface as *ħallī-(h)un* (let.IMP.2SG-3PL) as it would in Beirut/Damascus Arabic. Here again, the dialects of northern Syria and southern Turkey have borrowed the structure, but not the substance, and use the inherited root *mək* ‘let’, as exemplified in (21).

- (20) Beirut/Damascus Domari
 ħalli ġānd dfən lakrand-əs
 let go.SBJV.3PL bury do.SBJV.3PL-3SG
 ‘Let them go and bury him.’

- (21) Aleppo Domari
 mək pāvər pāsōr
 let come.SBJV.3SG 2SG.AD
 ‘Let him come to your place.’

3.3.5 Negation

Only two Arabic negators have made their way into the grammar of Northern Domari: Damascus Arabic *mū* and the contrastive negative coordination markers *lā...walā* ‘neither...nor’. Arabic *mū* is only available in the dialect of Beirut/Damascus. Its distribution and functions, however, only partially match those of Damascus Arabic. The primary function of *mū* in Damascus Arabic is to negate non-verbal predicates. This is not attested in Domari, which relies for this purpose only on inherited *nye*. *mū* surfaces first when negation has scope over non-clausal constituents, as shown in (22), and second when the predicate is in a non-indicative mood (subjunctive, jussive and imperative) as in (23):

- (22) Beirut/Damascus Domari
 səff (h)ra wāšya mū wāšōm
 side become.PFV.3SG 3PL.COM NEG 1SG.COM
 ‘He took sides with them, not with me.’

- (23) Beirut/Damascus Domari
 biğüz mū māntyar wāš məşrī
 possible NEG stay.SBJV.3SG 3SG.COM money
 ‘He might not have any money left.’

The Arabic structure *lā...walā* is readily available in all varieties, but whereas it is the only option in Beirut/Damascus, it competes with the inherited structure *nə...nə* in northern Syria and southern Turkey. Interestingly, this clash has led to a mixed form *nə...walā*, as shown in (24). The Domari syntax is also reminiscent of the Turkish possessive predication syntax with possessive marking on the noun and an existential morpheme.

- (24) Antioch Domari (Southern Turkey)
 nə lawr-ōs ašti wala šarš-ōs ašti
 NEG tree-3SG EXS NEG root-3SG EXS
 ‘It doesn’t grow on a tree nor has it roots.’

3.3.6 Complex sentences

Complex sentences minimally include coordinated and subordinate clauses. The Arabic coordinators *w* ‘and’, *aw* ‘or’, *walla* ‘or’, *bass* and others have all made their way into Domari. Originally, Domari seems to have distinguished clausal coordination from phrasal coordination, a not so frequent feature from a typological point of view. Nominal categories are coordinated with the Turkish-derived morpheme *la* and clauses are coordinated with the Kurdish-derived enclitic *-ši*. The intrusion of Arabic *w*, which in Arabic is used indiscriminately for both kinds of coordination, has led to the marginalization of the original system in Beirut/Damascus Domari, which now tends to favour the use of Arabic *w*.

- (25) Beirut/Damascus Domari
 illi mangar tōre māṣṭ-a-ma w illi mangar
 REL want.IPFV.3SG put.IPFV.3SG yoghurt-OBL-IN and REL want.IPFV.3SG
 ?ār-s-e nāšif
 eat.IPFV.3SG-OBJ.3SG-PRS dry
 ‘Some eat it in yoghurt, some eat it dry.’

As far as phrasal coordination is concerned, some alternation between Arabic *w* and Turkish-derived *la* is still observed: *dōmwārī u ṭāṭwārī* ‘Domari and Arabic’ ~ *dōm la šarabi* ‘Domari and Arabic’.

Virtually all the conjunctions of subordination found in Domari are borrowed from Arabic. This includes the relativizer *illi*, the complementizer *inno* and potentially all the adverbial conjunctions found in Levantine Arabic: *lamma* ‘when’, *qabəl-mā* ‘before’, *baʿəd-mā* ‘after’, *ša-bēn-mā* ‘by the time’, and many more. Pre-Arabic constructions are attested for relativization and conditional clauses,

Bruno Herin

but these only survive in the dialects of northern Syria and southern Turkey, and tend to be replaced by Arabic material (except in the varieties spoken in Turkey). A case in point is conditional clauses. Arabic *iza* and *law* are available everywhere, even in Turkey, as shown in (26), recorded in Antioch. In this example, the speaker uses the Arabic morpheme *aza* (< *iza*) in the first sentence of the utterance, and no overt marking in the protasis, making parataxis a possible means to express condition. As far as counterfactual conditions are concerned, it appears that the dialect of Beirut/Damascus is fully congruent with Arabic in having borrowed also the morpheme *kān* in both the protasis and the apodosis, as shown in (27). The dialects of northern Syria and southern Turkey exhibit a native strategy using subjunctive mood and past marking in the protasis and perfective and past marking in the apodosis. The two clauses are coordinated with the Kurdish derived enclitic *ši* (28).

(26) Antioch Domari

aza kām karne qāne kām nə-karne nə-qāne
if work do.IPFV.1PL eat.IPFV.1PL work NEG-do.IPFV.1PL NEG-eat.IPFV.1PL
'If we work, we eat, (if) we don't work, we don't eat.'

(27) Beirut/Damascus Domari

law kām nəčnār-sā bāb-ōm kām
if be.PRF.3SG make.dance.IPFV.3SG-OBJ.3PL father-1SG be.PRF.3SG
abšar kaki (h)re
not.know what become.PFV.3SG
'If my father had put them to dance, I don't know what would have happened.'

(28) Sarāqib Domari

all-əs byātyənd-a nə-ktēnd-s-a ši
God-ACC fear.SBJV.3PL-PST NEG-throw.PFV.3PL-OBJ.3SG-PST and
'Had they feared God, they would not have thrown him.'

3.4 Lexicon

3.4.1 Function words

Arabic prepositions do occur in Domari, but these are mostly non-core prepositions such as *qabəl* 'before', *bašad* 'after', *minšān* 'for', *yēr* 'other'. Some have made their way into Domari only recently, and still alternate with pre-Arabic options, such as the Iranian morpheme *war*, which tends to be replaced by Arabic

mitəl ‘as, like’ especially in the dialect of Beirut/Damascus. Currently, *war* and *mitəl* are in a quasi-complementary distribution with *war* being used with full NPs and *mitəl* with pronouns, as shown below in (29) and (30):

- (29) Beirut/Damascus Domari
 tō ʔr-ōm war ištōr
 you son-1SG like cop.2SG
 ‘You are like my son.’
- (30) Beirut/Damascus Domari
 tāni ʔər gēna mitl-ōs kām karre
 second son also like-3SG work do.IPFV.3SG
 ‘My second son has the same job.’

The Arabic core preposition *b-* ‘in, with’ occurs in Domari, but it appears to be restricted to certain constructions and idioms such as *gāl b-gāl* ‘discussion’ (word in-word), *ārāt əb-dīs* ‘night and day’ (night in-day), *b-rəbʃ-āk* ‘for a quarter of a pound’ (with-quarter-INDF). The preposition *min* ‘from’ also sporadically occurs in Beirut/Damascus Domari:

- (31) Beirut/Damascus Domari
 min ši šēš mās ġərsa krōm dāwat-ōs
 from about six month wedding-3SG do.PFV.1SG wedding-3SG
 ‘Some six months ago I married (my son).’

Domari also borrows high-frequency adverbs, fillers, connectors and all kinds of discourse-structuring devices, such as *masalan* ‘for instance’, *abadan* ‘at all, never’, *yaʃni* ‘I mean’, *aywa* ‘yes, so’, *waʃlla* ‘I swear’, *inno* (complementizer and discourse marker) and many more. One finds also common adverbial phrases such as *ṭul in-nhār* ‘all day long’, *ṭul il-waʔət* ‘all the time’, *ʃala ṭul* ‘immediately’. The very common Domari phrase *tika tika* ‘slowly’ replicates Arabic *šwayy ašwayy*.

3.4.2 Content words

In Syria and Lebanon, Arabic is the *de facto* lexical reservoir of Domari, so there is a general licence to integrate any element from Arabic if no pre-Arabic option exists. The issue is the replacement of pre-Arabic options with Arabic material. There is of course a certain amount of variation in lexical knowledge across speakers, but it seems possible to differentiate several levels of replaceability. Some

Bruno Herin

items have long been replaced by Arabic words, and only a handful of speakers are able to retrieve them, such as *lōrga* ‘tomato’ or *pīsanga* ‘bulgur’, replaced respectively by Arabic *bandōra* and *baryəl*. Other items tend to be replaced by Arabic equivalents but may still surface in the speech of some speakers, such as *čatīn* ‘hard’, *čirkī* ‘bird’, *alčāḥ* ‘low’ replaced by Arabic *šaṣab*, *ṭēr* ~ *šaṣfūr* and *wāṭi*. Some items seem stable but are sporadically replaced with Arabic-derived items such as *drəs kar-* ‘study’ instead of inherited *sək-*. Finally, other items such as *ḡawwəz h-* ‘get married’ and *ḡirsāwī h-* freely alternate. It appears therefore that every pre-Arabic item is somewhere on a continuum of replaceability from “very unlikely” to “completely disappeared”. To illustrate the variability in replaceability judgment, I remember an elicitation session in Aleppo with a father and his son. One of the sentences contained the Arabic word *baṣal* ‘onion’. The son simply translated the sentence with the Arabic word *baṣal* but the father strongly objected to this answer, stating that the proper Domari word is *pīwāz*.

As noted above, Arabic nouns are integrated in their singular form, except in the case of lexicalized plurals. Adjectives are borrowed in their masculine form and never agree in gender, as shown in (32). Other than the past copula *a*, all the words in this example are Arabic. Two features, however, allow its identification as Domari. First, *ḥāla* is realized without raising (also stressed on the last syllable [ħaːˈla]), unlike Levantine Arabic *ḥāle*, and second *taṣbān* does not agree in gender with *ḥāla* and surfaces in its masculine form, instead of feminine *taṣbāne*, as it would normally occur in Arabic.

- (32) Beirut/Damascus Domari
 ʔabəl ḥāla taṣbān a
 Before situation tired COP.PST
 ‘Before, the situation was bad’

Arabic verbs are easily integrated into Domari, because Domari has a light verb strategy. Roughly, transitive verbs tend to be integrated with the light verb *kar-* ‘do’: *rabbī kar-* ‘raise’ from Arabic *rabba*, *yrabbi* ‘raise’. Intransitive verbs are integrated with *h-* ‘become’: *ṣīš h-* ‘live’ from Arabic *ṣāš*, *yṣīš* ‘live’. While all the verbs that are integrated with *kar-* are transitive, some verbs integrated with *h-* are not intransitive: *lməs (h)rōs-s-e* ‘he has touched it’ (touch become.PFV.3SG-3SG-PRS) from Arabic *lamas*, *yilmis* ‘touch’. This seems to happen with transitive verbs that are lower on the transitivity scale, or at least perceived to be so. In the case of *lamas*, *yilmis*, its integration into Domari by way of the light verb *h-* suggests that speakers perceive it as less transitive. Formally, speakers isolate the imperfect stem of the verb, and apply a vocalism in /i/: *nsī kar-* ‘forget’ and

13 Northern Domari

stannī kar- ‘wait’, from the Arabic imperfect stems of *nsa* ‘forget’ and *stanna* ‘wait’.³ An exception to this tendency occurs with the so-called hollow roots in Arabic whose imperfect stem is CūC. In this case, speakers simply extract the imperfect stem and leave it unchanged: *zūr h-* ‘visit’, *dūr h-* ‘turn’, *ṣūz h-* ‘need’, from the Arabic imperfect stems *zūr*, *dūr* and *ṣūz*.

Some English-derived items were also recorded in the dialect of Beirut/Damascus, such as *mōmari* ‘memory card’, *hambarga* ‘hamburger’ and, more surprisingly, *tōmanǧīre* ‘Tom and Jerry’ [to:mandʒi:re], expectedly stressed on the last syllable.

3.4.3 Speech sample

Probably the best way to capture how Arabic integrates into Domari is to consider a piece of spontaneous speech, reproduced below in (33). It is part of a recorded discussion I had with a consultant in her mid-thirties in Beirut. It illustrates the level of endangerment of Beirut Domari. The consultant belongs to the last generation of fluent speakers. Her children did not acquire the language. According to what she reports, she was unable to speak to her children in their early childhood because her husband, who is a semi-speaker of Domari, prevented her from transmitting the language. Her daughter-in-law, aged 21 at that time, is also a fluent speaker of Domari because she grew up in Damascus, where language transmission was more solid than in Lebanon. Both of them use Domari in the household. Her son reacts negatively when he hears it, and even labels it *aǧnabi* ‘foreign, non-Arabic’. Linguistically, the text illustrates some of the features discussed above. Arabic-derived items are marked in boldface.

(33) Beirut Domari

nā n-ǧib karre pāṅǧī gāl karre gāl karre
 no NEG-tongue do.IPFV.3SG 3SG word do.IPFV.3SG word do.IPFV.3SG
 dōm wāšōm mā gāl kame wāšī **ṣadi** **bass** əʔr-ōm ʔzīn
 Dom 1SG.COM 1SG word do.IPFV.1SG 3SG.COM normal but son-1SG shout
 karre wat ftyare ma-gāl ka **aǧnabī**
 do.IPFV.3SG 3SG.SUPR say.IPFV.3SG NEG-word do.SBJV.2SG foreign
 nə-fəmm (h)ōme watōr, gāl karse
 NEG-understand become.IPFV.3SG 2SG.SUPR word do.IPFV.2PL
ṣarabiy-a-ma yaṣni ma-gāl k(a) **ēhāni laʔanno**
 Arabic-OBL-IN I.mean NEG-word do.SBJV.2SG so because

³These verbs are only available in Beirut/Damascus, other dialects use respectively *ziwra kar-* and *akī kar-*.

Bruno Herin

n-fəmm (h)ōre watī bass mā l pāngī ġib
 NEG-understand become.IPFV.3SG 3SG.SUPR but 1SG and 3SG tongue
 kane tūl il-waʔət kəry-a-ma yaʃni iza mā l pāngī
 do.IPFV.1PL length DEF-time house-OBL-IN I.mean if 1SG and 3SG
 štēn kəry-a-ma tūl in-nhār gāl kane dōm-a-ma
 COP.1PL house-OBL-IN length DEF-day word do.IPFV.1PL Dom-OBL-IN
 yaʃni ʔr-ōm wāri fəmr-ōs wāḥad u ʔišrīn sane akbar
 I.mean son-1SG bride age-3SG.F one and twenty year bigger
 ʔr-ōm-ki b-trən wars mū ʔādi ʔādi nye amīn lāzim
 son-1SG-ABL with-three year NEG normal normal COP.NEG 1PL must
 lpāran azyar wēšōma bass bxēz e u ādami e u
 take.SBJV.1PL smaller 1PL.ABL but good COP and humane COP and
 maḥšūm e mā ēhāny-a xr-a kē pārdōm-əs ʔr-ōm
 respectful COP so so-OBL heart-OBL BEN take.PFV.1SG-OBJ.3SG son-1SG
 kē u ġamāʃt-ēm kē skīr(a) ēta baʃdēn skīra
 BEN and folks-1SG BEN learn.PFV.3SG here then learn.PFV.3SG
 mahná baʃdēn kām əkra wars-ā wars-ā nīm makanīk
 profession then work do.PFV.3SG year-INDF year-INDF half mechanic
 baʃdēn wəndrārda u isa nə-kām kištar wala kkyā
 then fire.PFV.3SG and now NEG-work do.PROG.3SG nor thing
 wēsre kəry-a-ma
 sit.PFV.3SG house-OBL-IN

‘No, [my son] doesn’t speak [Domari], [my daughter-in-law] does, she speaks with me, I speak with her normally but my son shouts at her and tells her: “Don’t speak foreign, I don’t understand you, you all speak in Arabic, don’t speak like this”, because he doesn’t understand her. But me and her we speak all the time in Domari, that is, if both of us are in the house, all day long we speak in Domari. The bride of my son, she is 21 years old, three years older than my son, it’s not usual, we [women] have to take someone older, but she is a good person, humane and respectful. That’s why I took her for my son and my family. [My son] studied here [in the school]. After that he went for vocational training and worked for a year a year and a half as a mechanic – then he renounced. And now he doesn’t do anything, he stays home.’

4 Conclusion

Multilingualism seems to have been a normal state of affairs amongst the Doms for a very long time, probably since the genesis of the community. The reason for this is mostly because the sociolinguistics of Domari has in likelihood remained unchanged throughout the centuries: Domari is a community language whose use is restricted to in-group communication. Out-group interactions imply the use of the majority language. Due to the very nature of their occupational profile, peripatetic groups are forced to have frequent interactions with outsiders. This involves *de facto* high levels of bilingualism. Although it is hard to assess whether the dominant language is the insider code or the outsider code, it makes sense to suspect that balanced bilingualism was the norm, as much in the past as in the present.

Van Coetsem (1988; 2000) uses the term “transfer” generically for any kind of contact-induced phenomenon. If the transfer is triggered by speakers who are dominant in the source language, he uses the term “imposition”. If it originates from recipient-language dominance, it is called “borrowing”. Lucas (2015: 525) further introduces two categories, the first of which he calls “restructuring”, defined as a “type of change [...] brought about by speakers for whom the changing language is an L2, but it does not involve transfer”. He notes that for individuals who acquired two languages simultaneously (in early childhood), “the distinction between borrowing and imposition breaks down”. In this case, both languages typically undergo “convergence”, that is the fourth category of contact-induced change. Because I posit balanced Arabic–Domari bilingualism as the norm, the question that needs to be answered is whether all the contact-induced changes happening in Domari are the product of convergence, or whether there are changes that can be attributed to Arabic dominance (source-language agentivity or imposition). Another problem concerns the sociolinguistic limits of the model. Speakers with two first languages are expected to initiate changes that target both languages. When languages exhibit unbalanced sociolinguistic statuses (minority versus majority), one wonders how changes originating from minority language agentivity can diffuse to the majority. Although it cannot be ruled out, it remains very unlikely. Consequently, convergence will always happen in the direction of the minority language. And this is indeed what is happening between Arabic and Domari: they become more and more similar at all levels, but only Domari is moving towards Arabic.

In the realm of phonology, it was shown that Domari has kept a distinct inventory from Arabic, although convergence with Arabic is almost complete for

Bruno Herin

short vowels. A possible consonantal imposition is found in Beirut/Damascus Domari where etymological /q/ is realised as /ʔ/, as in neighbouring Arabic dialects. As far as morphology is concerned, eligible candidates for imposition are the Kurdish diminutive *-ək*, the Turkish conditional clitic *sa* and superlative *ān*. An evident case of imposition is the phenomenon that seems the most sensitive to dominance: so-called “bilingual suppletion” (Matras 2012). Bilingual suppletion in Northern Domari can be observed only in the dialect of Beirut/Damascus in the case of comparatives, and incipiently in the case of numerals. As far as syntax is concerned, cases of imposition are probably the transfer of Arabic auxiliaries and the negator *mū*. The transfer of utterance modifiers such as fillers, adverbs, conjunctions and virtually all discourse structuring devices are so prone to replication in contact situations (Matras 1998) that it is difficult to assess the source of agentivity. Other features discussed in this paper, such as constituent order, the internal object and the impersonal construction are clear instances of convergence.

As noted above, the main direction of change in Domari is towards convergence with Arabic, as expected in case of absence of dominance. The Beirut/Damascus dialect is the most convergent of all the Northern dialects, which in itself suggests that Arabic–Domari bilingualism is older in that variety. The Arabic component in Domari is largely uneven cross-dialectally and no overall statement about its nature can be made. The general picture that arises is that the impact of Arabic gradually increases from north to south, with the dialects located in northern Syria and southern Turkey being the least arabized, the Southern dialects spoken in Palestine and Jordan being the most influenced by Arabic, and the dialect of Beirut/Damascus exhibiting an intermediary stage. It was also shown that the main difference between Northern and Southern Domari as far as Arabic is concerned is the reluctance in Northern Domari to transfer Arabic inflections and the general tendency to favour the transfer of structures without substance.

Further reading

For a general account of the Arabic component in all the varieties of Domari documented so far, see Herin (2018). The paper discusses the Arabic component in Southern and Northern dialects. This is the only paper that tackles extensively the issue of contact-induced change in Domari from a global perspective.

For a description of the Domari dialect of Aleppo, readers can refer to Herin (2012).

Herin (2014) identifies the grammatical features that make Northern Domari a

coherent dialectal group.

Herin (2016) investigates the full extent of variation in Domari as a whole, drawing on data from both Northern and Southern Domari.

Readers can refer to Matras (this volume) for a list of references about Jerusalem Domari.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	INDF	indefinite
ABL	ablative	IPFV	imperfective
ACC	accusative	M	masculine
AD	adessive	NEG	negation
BEN	benefactive	NP	noun phrase
CMPR	comparative	OBJ	object
COM	comitative	OBL	oblique
COP	copula	PFV	perfective
DEF	definite article	PL	plural
DEM	demonstrative	PRS	present
EXS	existential	PRF	perfect
F	feminine	PROG	progressive
IMPF	imperfect	PST	past
IPA	International Phonetic Alphabet	REL	relative
IN	inessive	SBJV	subjunctive
IND	indicative	SG	singular
		SUPR	superessive

References

Herin, Bruno. 2012. The domari language of aleppo (Syria). *Linguistic Discovery* 10. 1–52.

Herin, Bruno. 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.

Herin, Bruno. 2018. The Arabic component in domari. In Stefano Manfredi & Mauro Tosco (eds.), *Arabic in contact*, 19–36. Amsterdam: John Benjamins.

Bruno Herin

- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Macalister, Robert Alexander Stewart. 1914. *The language of the Nawar or Zutt, the nomad smiths of Palestine*. Edinburgh: Edinburgh University Press.
- Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36(2). 281–331.
- Matras, Yaron. 2012. *A grammar of Domari*. Berlin: De Gruyter Mouton.
- Stassen, Leon. 2013. Comparative constructions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/121>.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.

Chapter 14

Mediterranean Lingua Franca

Joanna Nolan

SOAS University of London

This chapter explores the effect of Arabic contact on Lingua Franca, an almost exclusively oral pidgin spoken across the Mediterranean and along the North African coastline from the seventeenth–nineteenth centuries. The chapter highlights the phonological and lexical impact Arabic appears to have had on Lingua Franca.

1 Overview and historical development

Today, a LINGUA FRANCA is a term describing a language used by two or more linguistic groups as a means of communication, often for economic motives. Typically, none of the groups speak the chosen language as their native tongue. The original and eponymous Lingua Franca, however, was a trading language, used among and between Europeans and Arabs across the Mediterranean (Kahane & Kahane 1976). Its exact historical and geographical roots – as well as its precise lexifier languages – prove elusive. Hall (1966) dates Lingua Franca’s birth to the era of the crusades, while other linguists (Cifoletti 2004; Minervini 1996) suggest that it took root on the North African Barbary Coast (in the Regencies of Algiers, Tunis and Tripoli) at the close of the sixteenth century.

Contention extends to its very name. There are several discrete etymological suggestions for the term *Franca*. Some linguists interpret *Franca* as meaning ‘French’ (e.g. Hall 1966: 3). Hall claims that France’s regional significance in the medieval era meant that its languages, specifically Provençal, were adopted across the Mediterranean, and were a key constituent of the original Lingua Franca. In their etymological study of Lingua Franca, Kahane & Kahane (1976: 25), on the other hand, assert that the name *Lingua Franca* is rooted in the East and the Byzantine tradition, stemming from the Greek word *phrangika*, which



Joanna Nolan

denoted Venetian as much as Italian, or indeed, as any Western language (Kahane & Kahane 1976: 31). An alternative etymology for *Lingua Franca*, espoused by Schuchardt (1909: 74) among others, is from the Arabic, *lisān al-faraṅġ* ‘language of the Franks’. This initially referred to Latin and then to describe a trading language employed largely by Jews across the Mediterranean. It later came to encompass the languages of all Europeans, but particularly Italians (Kahane & Kahane 1976: 26).

Evolving from its maritime origins, by the late sixteenth century, *Lingua Franca* was the language of pirates of the North African Barbary coast and their captured slaves, and, as such, the subject of legend and myth. Indeed, the variation found in the accounts of *Lingua Franca*, and descriptions of its linguistic makeup lead some linguists (Minervini 1996; Mori 2016) to suggest that there may have been multiple *Lingua Francas* or that it was simply L2 Italian. As Schuchardt (1909: 88), identified, *Lingua Franca* – perhaps above all in its resistance to theoretical classification – adheres to Heraclitus’ philosophy of *panta rei* ‘everything is in flux’.

Contemporaneous descriptions of *Lingua Franca* detailing its lexifiers and, in some cases, its salient features, come mostly from the North African Barbary Regencies and from the Levant. While the writers of these descriptions often identify Italian and Spanish as lexifiers, there are also, if fewer, mentions of Portuguese, French, Provençal, Arabic, Turkish and Greek (see below for further detail). This speaks to the hypothesis that there were multiple *Lingua Francas*, or perhaps more appropriately *lingua francas*. It also raises the frequent subjectivity of the source’s writer and their consequent interpretation of *Lingua Franca*. Their native language appears to have a bearing on the makeup of the *Lingua Franca* recorded. It may influence the lexicon they hear, as well as the orthography they employ in their account. Equally, there is the subjectivity of the researcher to bear in mind. The assumption that a French source, for example, has represented *Lingua Franca* in a particular manner overlooks the fact that the European residents, particularly of port cities across the Mediterranean, most likely would have been multilingual, with an ability to adapt their lexicon to maximize understanding and communication with their interlocutor.

The most widespread documentation of *Lingua Franca* comes from the Levant and into northwest Africa. Algiers, and to a lesser extent, Tunis and Tripoli, had long been the crucible of Mediterranean piracy, and as the slave trade of Barbary pirates increased – with over a million European slaves held there between the sixteenth and nineteenth centuries (Davis 2004: 23) – so too did the domains and usage of *Lingua Franca*. The sixteenth–seventeenth century Spanish Abbot

Diego del Haedo described it as follows:

La que los Moros e Turcos llaman Franca... siendo todo una mexcla de lenguas cristianas y de vocablos, que son por la mayor parte Italianos e espanoles y algunos portugueses...Este hablar Franco es tan general que non hay casa do no se use

‘that which the Arabs and Turks call Franco...being a mix of Christian languages and words, which are in the majority Italian and Spanish and some Portuguese, this Franco is so widespread that there isn’t a house [in Algiers] where it isn’t spoken’ (Haedo 1612: 24; author’s translation).

Despite its alleged profusion in the Barbary coast, and numerous references in various contemporary sources, the corpus of Lingua Franca is remarkably limited. The exclusively European documentary sources (from contemporaneous diplomats, travellers, priests and slaves) provide mostly phrases and individual words and a handful of short dialogues. The most fulsome examples come from literature, and, as such, only provide indirect, and potentially less authentic, evidence of the contact vernacular.

For example, the alleged earliest record of Lingua Franca comes from an anonymous poem, *Contrasto della Zerbitana*, found by Grion (1890) in a fourteenth-century Florentine codex, and apparently written in the late thirteenth or early fourteenth century on the island of Djerba, off the coast of Tunisia. The sixteenth century *La Zingana* (Giancarli 1545) has an eponymous Arabic-speaking heroine whose language features hallmarks of Lingua Franca, while the speeches of the Turkish characters in Molière’s *Le Sicilien* (1667) and *Le Bourgeois gentilhomme* (1798 [1670]) also appear to share a number of its defining linguistic traits.

The first detailed description and documentation of Lingua Franca comes in Haedo’s (1612) *Topographia*, a comprehensive study of Algiers, with a chapter devoted to the languages spoken there. Haedo spent several years in Algiers at the close of the sixteenth century and was even imprisoned for a number of months. His *Topographia* details the urban features, social makeup and linguistic mix of Algiers, creating an impression of Lingua Franca’s ubiquity across multiple domains, and indispensability to daily commercial, and even domestic, life.

Other early sources are predominantly French. A Trinitarian priest, Pierre Dan, was almost contemporary with Haedo in Algiers, while in the mid-seventeenth century the diplomat Savary de Brèves travelled to Tripoli, and Chevalier D’Arvieux, King Louis XIV’s envoy to the region, and advisor to Molière on Turkish and Arabic matters, visited both Algiers and Tunis. All these men offered excerpts

Joanna Nolan

of Lingua Franca in their writings, as well as descriptions of its character and lexifiers (Savary de Brèves 1628; Dan 1637; D'Arvieux 1735).

Certainly seventeenth-century Algiers and the other two Barbary Regencies of Tunis and Tripoli provided the conditions for what had previously been a prepidgin – with limited lexicon and a lack of stability – to evolve into the language of daily life, permeating all echelons of society and facilitating contact among and between the plurilingual populations of the urban centres. The Barbary states were, from the late sixteenth century, under the *de jure* but not *de facto* control of the Ottoman empire, whose support was needed to shore up the rule of the Greek Barbarossa brothers who had ousted Spanish forces from North Africa. The two brothers (named for their red beards), Aruj and Hizir, gradually brought much of North Africa under Turkish sovereignty through a series of naval challenges and, later, city sieges securing power over coastal areas. The indigenous population rallied to the brothers' cry and although the elder, Aruj, was slain, Hizir assumed control of Algiers in the early sixteenth century (Tinniswood 2010: 8; Weiss 2011: 10). He immediately offered the Ottoman Empire control over the brothers' conquests in order to bolster his own position and ward off threats from Spain. Ottoman rule was compounded over the following decades (Clissold 1977: 27).

In reality, however, the Regencies had unstable political systems with local elites vying for power. Their economy was driven by corsairing, and the real power lay in the hands of the mostly European renegades who carried out raids on land and at sea, seizing cargo and most importantly human booty, sold as slaves on their return (Plantet 1889).

The huge influx of captured Europeans swelled the urban population and created multinational, multid denominational and notably multilingual societies. The Flemish diplomat D'Aranda, imprisoned in Algiers in the 1660s, wrote of hearing 22 languages in the slave quarters of the city (D'Aranda 1662). Lingua Franca emerged as a contact language accessible to the majority of slaves (though not all), given its Romance-influenced lexicon. Although the elites were predominantly Arabic speaking, Europeans permeated the upper levels of Barbary society through their economic sway as corsairs and high levels of inter-marriage of Arabs and Europeans. Lingua Franca quickly became the default language within the slave quarters, known locally as *bagnios*, seemingly a Lingua Franca term, and in master–slave relationships. Authors who detail the use of Lingua Franca across more than 250 years and throughout the regencies, including Pananti (1841) and Broughton (1839), report the regular use of Lingua Franca by Arabic-speaking slave owners, including the Pashas, Beys and various dignitaries of the ruling

households.

As noted above, Lingua Franca also elicits various opinions regarding its key lexifiers, though one common point of agreement among its contemporary witnesses and speakers is that Italian¹ and Spanish are mentioned repeatedly as principal sources. Descriptions mostly include at least three lexifying languages, though not always the same three: while Italian and Spanish are consistently named, Provençal also features (D'Arvieux 1735: vol. 5, p.235) as does Portuguese (D'Aranda 1662: 22; Haedo 1612: 24). A much later account from the Italian merchant Pananti, briefly imprisoned in Algiers, mentions Arabic as a lexifier: *la Lingua franca è una mistura d'italiano, di arabo e di spagnolo* 'Lingua Franca is a mix of Italian, Arabic and Spanish' (Pananti 1841: 201). Within the same memoir, however, Pananti refers to African rather than Arabic as one of three lexifiers of Lingua Franca (the other two being Italian and Spanish). Such inconsistency is the hallmark of many of the sources, compounding an already confused and often contradictory picture of the language. As Selbach (2008: 18) observes, "lexical variants were as much a part of the language as variant lexifiers."

The contribution of multiple languages to Lingua Franca is borne out by its lexicon: there are often several alternatives for a single meaning, as listed in the *Dictionnaire* (Anonymous 1830), the sole comprehensive lexical record of the language. For example, 'to do' is translated by *far* (from the Italian *fare*), *fazir* (from Portuguese *fazer*), and *couchiar* (likely from Sicilian *cunzari*; Cifoletti 2004: 316). Lingua Franca – as substantiated by its corpus of written attestations – is always labelled as such by a non-speaker. The European authors of descriptions of Lingua Franca, its diffusion and its usage, present it as, if not foreign, certainly removed and remote from their own languages, and attribute the speaking of Lingua Franca to the Arabs and Turks (even if it is clearly a language that lexically is much closer to European, and specifically Romance, languages).² In her comprehensive if subjective account and an analysis of Lingua Franca, Dakhliia highlights how citations, often expressing insults and aggressive warnings, and usually introduced into the texts by witnesses to Lingua Franca in direct speech, punctuated with exclamation marks, underline what she terms *le choc linguistique de l'altérité et de la barbarie* 'the linguistic shock (or jolt) of otherness and barbarism' (Dakhliia 2008: 351). This choice of words is important. Dakhliia's asso-

¹Italian is a catch-all term, used by contemporary authors in Barbary, as well as linguists today, as identified by Trivellato (2009: 178): "I write 'Italian', 'Portuguese' and 'Spanish', but recall that European written languages in the epoch were not fully standardized." Venetian and Tuscan were both described as Italian, for example.

²See, for example, the exchange between Louis Bonaparte and Hyde Clarke across several issues of the periodical *Athenaeum* in 1877.

Joanna Nolan

ciation of otherness and barbarism suggests that Lingua Franca, according to the European documentary sources, was the language of the Arab oppressor. While this may apply to the corsairs and slave-masters, speaking Lingua Franca to their European captives, there are other instances within the corpus of written attestations where Arabic-speaking elites use Lingua Franca in diplomatic, even philosophical, exchanges. For example, Louis Frank, the Bey of Tunis' doctor, comments on the deemed impropriety of the Bey speaking formal Italian, and his consequent use of Lingua Franca, which permeated all levels of society (Frank 1850: 70):

la langue franque, c'est à dire cet italien ou provençal corrompu qu'on parle dans le Levant, lui est également familière; il avait même voulu essayer d'apprendre à lire et à écrire l'italien pur-toscan: mais les chefs de la religion l'ont détourné de cette étude, qu'ils prétendaient être indigne d'un prince musulman.

'lingua franca, or rather this bad Italian or Provençal spoken in the Levant, is equally familiar to him; he had actually wanted to learn to read and write pure Tuscan Italian; but his religious chiefs had warned him off such study, which they claimed was unworthy of a Muslim prince' (author's translation).

Frank wrote of the intriguing linguistic, socio-political, cultural and even religious conflation evidenced in Lingua Franca, describing an encounter with a Muslim beggar, who implored: "*Donar mi meschino la carità d'una carrouba³ per l'amor della Santissima Trinità e dello gran Bonaparte*", 'Please to give miserable me the charity of a penny for the love of the most holy Trinity and the great Bonaparte' (Frank 1850: 101; author's translation). In just this one Lingua Franca sentence, multiple lexifiers are represented: *meschino* is from the Arabic, *miskīn*, and there is Spanish in *carrouba* 'penny', *donar* 'give', and *amor* 'love', with the Italian Catholic reference of *Santissima Trinità* 'most holy Trinity', and French *Bonaparte*. The latter would have still been Emperor and possibly at the height of his power. Bonaparte is qualified by *gran* 'great', from the Italian or even Venetian. It suggests how cosmopolitan, multicultural and multilingual Tunis and its population had become that a beggar should speak this way. Even Frank was struck by the incongruity of the beggar's words: "*sa supplique en ces termes, bien étranges dans la bouche d'un Musulman*", 'his petition in these terms, very odd in the mouth of a Muslim' (Frank 1850: 101; author's translation).

³Until 1891 a *carrouba* was worth 1/16 of a Tunisian *piaster*, according to Rossetti (1999).

Lingua Franca's demise dates from 1830, as a consequence of the outlawing of slavery and the start of the French colonization of North Africa. Lingua Franca became known alternatively as Sabir (HSA 1882: Letter 6-7473), with a later incarnation which Schuchardt dubbed Judeo-Sabir (Schuchardt 1909: 87). Residual elements seem to persist, however, in other contemporary jargons and languages. The pidgin spoken in Algeria, Pataouète, while largely lexified by French and Arabic, also features a significant number of words that are identified by Lanly as Lingua Franca in origin. For example, Duclos' (1992) Pataouète dictionary enumerates at least 30 words whose etymology she specifies as Lingua Franca. These include *baroufa* 'quarrel', *fantasia* 'pride, delusion', *mercanti* 'merchant', and *ra-bia* 'rage'.

2 Contact with Arabic

As mentioned above, a substantial proportion of Lingua Franca speakers appear to have had Arabic as their first language. Inevitably, there would have been transfer when they spoke Lingua Franca, although, given the shared history of Arabic and European cultures in Sicily, Spain and other parts of the Mediterranean, it is perhaps hard to state unequivocally whether lexical influences stem from the contact of Arabic with Lingua Franca directly, or its earlier contact with Romance languages. Pellegrini (1972) identified the many Arabic loanwords integrated into Italian, particularly in the realms of trade, conflict and exploration. A number of these are included in the *Dictionnaire* (1830), including *magazino* 'shop' from the Arabic, *maḥāzīn* 'storage facility', and *fondaco* 'trading post' from the Arabic *funduq* 'hotel, inn'. Both would already have been in use in Italian, thus complicating further an etymological study of Lingua Franca's lexicon.

3 Contact-induced changes in Lingua Franca

Contact-induced changes in Lingua Franca with regard to Arabic are relatively limited, evident predominantly in its lexicon but also, to some extent, in its phonology. It is perhaps even an overstatement to consider Arabic's influence as contact-induced change; rather it might be viewed simply as an additional lexifier.

3.1 Phonology

The relative lack of written record and potential unreliability of the sources' excerpts of Lingua Franca makes the identification of a definitive phonemic inven-

Joanna Nolan

tory both difficult, and at times, inconclusive. Overall, Lingua Franca follows the phonology of Romance languages, predominantly Tuscan Italian, though with elements of Venetian and Spanish. Venetian influence is also evident in the Lingua Franca tendency to omit final vowels following sonorants /l/, /n/ and /r/, as in *colazion* instead of *colazione* ‘breakfast’. Both Venetian and Lingua Franca exhibit examples of degemination (e.g. *tuto* ‘all’ vs. Tuscan *tutto*) and voicing of intervocalic stops – *segredo* ‘secret’ rather than *segreto* (Ursini 2011). The voicing of intervocalic *t is also consistent with Spanish, which appears also to have had an influence on elements of Lingua Franca phonology. An example from the *Dictionnaire* (1830: 63) that illustrates both the plosive voicing and the final vowel omission is *padron* ‘master’, an epithet that recurs throughout the corpus of attestations. However, in terms of the language’s vocalic system, Arabic appears to exert some influence. Cifoletti (2004) suggests that Arabic influence on the realization of vocalic elements can be seen in the *Dictionnaire*: *bonou* from the Italian *buono* ‘good’ evidences a simplification of the diphthong *uo*. Camus Bergareche (1993: 444) confirms this simplification; for example Italian *uovo* ‘egg’, *duole* ‘hurt’, *buono* ‘good’ are reduced in their Lingua Franca counterparts: *obo*, *dole*, *bono*.

There is some evidence in the *Dictionnaire* of a reduction in the number of qualities for short vowels in Lingua Franca from a typical five-vowel Romance system to the more impoverished systems found in North African Arabic varieties, seen, for example in the frequent realization of final /e/ as <a>, as in *scoura* from *scure* ‘axe’ or *gratzia* from *grazie* ‘thanks’, or <i>, as in *sempri* from *sempre* ‘always’ or *grandi* from *grande* ‘big’. Camus Bergareche (1993: 444) reinforces this, citing the Lingua Franca words *mouchou* ‘much, many’, *poudir*, ‘to be able’, and *inglis* ‘English’, with their roots in Spanish (*mucho*, *poder*, *ingles*) as evidence of a reduction in vowel qualities as a result of contact with Arabic.

Unlike most (non-sonorant-final) words in Lingua Franca that have a typical Romance vowel ending, Arabic-derived words generally retain their consonant ending, as in *rouss* from *ruzz* ‘rice’ and *maboul* from *mahbūl* ‘stupid’ (Cifoletti 2004: 38). Non-Romance influence on Lingua Franca is also evidenced in the regular substitution of /b/ for /p/, which is lacking in the phonemic inventory of most Arabic varieties. The *Dictionnaire* features a number of replacements of this type (Cifoletti 2004: 38), as in *esbinac* ‘spinach’ and *nabolitan* ‘Neapolitan’. Minervini (1996: 257–60) analyses the speech of the eponymous heroine of Giancarli’s (1545) *La Zingana*, and comments on the frequent substitution of /b/ for /p/ and /v/, offering examples such as *cattiba* (*cattiva* ‘nasty’ in Italian), *bericola* (*pericolo* ‘danger’ in Italian), the native Arabic of the character allegedly influenc-

ing her pronunciation. Given, however, that these examples come from a work of fiction, they do not provide conclusive evidence of influence.

Perhaps given that *Lingua Franca* as attested is replete with abbreviation, elipsis and omissions, it predictably features examples of aphaeresis. For example, many Romance-derived items beginning with a syllable that resembles the Arabic definite article see this omitted in *Lingua Franca*. Examples include *bassador* for *ambasciatore* ‘ambassador’, *bastantza* for *abbastanza* ‘enough’, and *rigar* for *irrigare* ‘to water’. As with many other linguistic features, however, the similarity between *Lingua Franca* and Venetian dialect must be considered, as some of these words exist in a similarly abbreviated form in Venetian (Schuchardt 1909).

3.2 Lexicon

From my quantitative analysis of the material in the *Dictionnaire* and other available sources, it is apparent that there were very few Arabic lexemes in *Lingua Franca*’s lexicon. Of the more than 2,100 entries in the *Dictionnaire*, 32 are of Arabic origin, and of the 176 additional *Lingua Franca* lexemes identified in the corpus of attestations, only nine have an Arabic etymology. However, a number of the individual items are regularly repeated in the corpus, and, as such, Arabic appears a more influential lexifier on a token than on a type basis.

Romance/non-Romance (often Arabic) doublets feature particularly in terms of place names, and officialdom within the Regencies. The port of Tunis was known by its French, Italian and Arabic names, seemingly interchangeably: *La Goulette*, *La Goletta*, and *Wādi l-Ḥalq* ‘the gullet’. In his manual for future consuls, the outgoing English consul of Tripoli, Knecht, enumerates the hierarchies within the Pasha’s household and city administration. Many of these involve a combination of Arabic or Turkish and Italian, or perhaps *Lingua Franca*. Key positions include a *hasnadawr Grande ed un hasnadawr Piccolo* – ‘a senior treasurer and a junior treasurer’ (Pennell 1982: 97; author’s translation). *Hasnadawr* comes from the Ottoman Turkish *hazinedar* or *haznadar* (from Arabic *ḥāzin ad-dār*) ‘Lord treasurer of the household’ (Gilson 1987).

Another example is *Kecchia Grande* and *Kecchia Piccolo* ‘chief administrator and assistant administrator’ (Pennell 1982: 104). In the letters written by members of the household of Richard Tully, British Consul to Tripoli at the close of the eighteenth century, there is a reference to the “Great Chiah and the Little Chiah” (Tully 1819: 70), surely an anglicization of the title. *Kecchia* appears to derive from the Tunisian Arabic *kāhiya* ‘chief officer of an administrative district’ – *kecchia* is an Italianised (or, again, possibly *Lingua Franca*-influenced) orthography and pronunciation. Similarly, *sotto rais* (from the Italian and Arabic literally meaning

Joanna Nolan

‘under captain’) denoted the second in command of the harbour (Pennell 1982: 97, 100). The commander is referred to separately as the *rays de la marina* ‘chief of the port’ (Pennell 1982: 92). Again, one finds the combination of Arabic and Italian. (*Rays* is spelt in two different ways,⁴ which highlights how, pre-standardisation of European languages, orthography was erratic, even within a single document.) Another example is a proverb regarding the eradication of plague by the Day of St. John. Two variations on the proverb are cited – by Poiret (1802) and Rehbinder (1800):

- (1) (Poiret 1802)
Saint Jean venir, Gandouf andar
 ‘[The day of] St. John comes, the plague leaves.’
- (2) (Rehbinder 1800)
Saint Jean venir, buba andar
 ‘[The day of] St. John comes, the plague leaves.’

Selbach (2008: 44) remarks on how such varied nomenclature in Lingua Franca “allowed for much room to manoeuvre, and for speakers to mark their religious, political and cultural identity”. *Buba* ‘plague’ appears to come originally from the Greek, *βουβών* (*boubōn*) ‘groin’, suggesting yet another potential lexifying influence on Lingua Franca, while *Gandouf* plausibly derives from Arabic *yunduba/yundūb* ‘swollen tonsils’ (Schuchardt 1909: 72; cf. Selbach 2008: 45).

This example raises the issue of words already common to Arabic and Romance languages, since contact between them, as discussed above, had been prolonged and extensive. Similarly, the French (and possibly Italian or even Lingua Franca) word *avanie* ‘fine, insult, affront’ occurs often in the corpus of attestations (Pananti 1841; Grandchamp 1920). Grandchamp (1920: xiii) defines it thus:

les avanies étaient des sommes d’argent que les pachas réclamaient aux marchands des échelles sous les prétextes les plus divers, prétextes la plupart du temps injustes, parfois extrêmement bizarres

‘the fines were sums of money the Pashas demanded of the Levant merchants on various pretexts, pretexts that were for the most part unfair, and at times extremely strange’ (author’s translation).

⁴The standard Arabic form is *raʿīs*.

14 *Mediterranean Lingua Franca*

Although this word would appear to be derived from French, or at least a Romance language, given that it was the creation of Ottoman elites, it seems more likely that its origins are Turkish. This is confirmed by Pihan (1847) who suggests that it actually originally derives from the Arabic *hawān* ‘contempt’,⁵ but who also states (Pihan 1847: 46):

se dit également des impôts énormes que les Turcs font peser sur les Chrétiens dans le but de les humilier

‘it applies equally to the enormous taxes the Turks impose on Christians with the goal of humiliating them’ (author’s translation).

Additionally, there are words that appear to have etymologies in multiple languages that are rarely translated, at least by English sources such as Tully (1819). Two terms with similar meanings, *firman* and *teschera* ‘pass, decree’ and ‘pass, edict’, both issued by Ottoman or Arabic rulers, also bear remarkable similarity to Italian words with comparable meanings. *Firman* is from Arabic: *firmān* / *faramān*, though originally Persian, and would have come into Arabic through Ottoman Turkish, once again reinforcing how the languages spoken in the region were far from discrete entities. However, *firmare* in Italian means to sign, and the Lingua Franca translation of the French *seing* ‘signature’ and *signature* ‘signature’ in the *Dictionnaire* is *firmar*. A decree or pass (allowing free passage or safe conduct) would necessarily require an official signature. *Teschera* ‘pass, edict’ might appear to come from the Italian *tessera* ‘pass, ticket’ but there is also the Arabic word *taḍkira/taḍkara/tazkira/tazakara* (all variant realizations of the same item), which also means ‘permit’ or ‘ticket’. Both words seem integral to Barbary life, and are not translated. Tully (1819: 258) writes:

It is still affirmed that he has a *teskerra*, or *firman*, with him for this unfortunate Bashaw. A *teskerra* is a written order from the Grand Signior, and is held so sacred that every Musulman who receives it must obey its mandate, even to death.

Perhaps the most iconic word in Lingua Franca is *fantasia*. It is mentioned by multiple sources spanning more than two centuries (e.g. Haedo 1612; Broughton 1839). Although it appears to be a Romance word, Schuchardt (1909: 71) points out that it is used in the Arabic sense of pride, arrogance, as in, for example, Egyptian Arabic *itfanṭaz* ‘to give oneself airs’.

⁵This etymology is also favoured by *Le Trésor de la langue française informatisé* (Dendien 1994).

Joanna Nolan

Collections in the UK National Archives provide some limited evidence of borrowings from Lingua Franca (as opposed to other Romance languages) into Arabic. Hopkins' (1982) research in these archives focuses on two sets of English (and later British) state papers relating to the Barbary regencies and Morocco, including correspondence in Arabic, from the late sixteenth to late eighteenth centuries.

Hopkins adds a glossary to his translations that demonstrates the extent to which the Arabic letters from the Barbary States disproportionately feature loans from Romance languages. Hopkins (1982: x) comments that “[f]oreign words are very common and seem to be used quite unselfconsciously”, and many of those words Hopkins isolates are also listed in the *Dictionnaire* (1830) as Lingua Franca words. Three such items – *justisiya* ‘justice’, *markānti* ‘merchant’ and *zabantut* (*sbendout* in Lingua Franca, presumably from the Italian *bandito*) ‘pirate’ – occur in a single letter of unspecified provenance but written to the King of England in 1730 by a man claiming to be an Algerine trader in Tripoli (TNA: SP 71/23/51). The incidence of three non-Arabic, Romance words, is noteworthy and it seems plausible that these were Lingua Franca terms in such common usage that they would be borrowed as a native language alternative, an example of *mot juste* switching (Gardner-Chloros 2009: 32).

4 Conclusion

As demonstrated, the available corpus of writing in Lingua Franca – both documentary texts written by European visitors to the Barbary states and the dramatic works produced by contemporary authors – offers limited evidence of both lexicon and grammar. This makes description of Lingua Franca challenging, and, likewise, any concrete and substantiated analysis of its relationship with other, particularly non-European, languages.

Nevertheless, this chapter has suggested how Arabic and Romance languages influenced the emergence of Lingua Franca, specifically in terms of its lexicon and phonology. Authors throughout the era of Lingua Franca's existence, from Haedo (1612) to Broughton (1839) and Frank (1850) reiterate that, despite its overwhelming Romance base, Lingua Franca was spoken predominantly by the often Arabic-speaking slavemasters and rulers of the Barbary states. The plurilingual character of the population of this region, both collectively and individually, compounds an already unclear picture, however, as the fluidity of Barbary society led to European (and often Romance-language speaking) corsairs and diplomats alike permeating its upper echelons (Haedo 1612: 9; Garcès 2011: 129).

The lexical influence of Arabic is most evident in the Romance/Arabic doublets used in official terms and place names. These are often compound terms, such as *ra'is de la marina* ‘captain of the port’. Warrington, English Consul to Tripoli in the late eighteenth century, uses an Anglicised version of the phrase, *rays marina*, suggesting the ubiquity of such doublets (TNA: FO 161/9). Further evidence from the National Archives (Hopkins 1982) demonstrates that Lingua Franca words were borrowed in correspondence from Arabic-speaking dignitaries in the Barbary Regencies to the English Secretary of State, showing evidence of Lingua Franca borrowings in the written as well as the oral domain.

In terms of phonology, there seems to be evidence offered by the *Dictionnaire* (1830) and the observation of Haedo (1612) of the influence of Arabic on Lingua Franca. Haedo (1612: 24) also stated, with regard to the “Moors and Turks” that “*no saben ellos variar los modos, tiempos y casos*”, ‘they don’t know about gender, tenses and cases’ (author’s translation). Given that Lingua Franca lacks for the most part any verbal inflection and an absence of cases, this might be, as Haedo says, a result of contact, but it is typical of most pidgins and cannot be attributed solely to contact with Arabic. Such lack of certainty applies more generally. Arabic evidently exerted some influence on the evolution of Lingua Franca in North Africa, but not to the extent that it can straightforwardly be classified as contact-induced change.

Further reading

Nolan (2018) gives a comprehensive English-language introduction to Lingua Franca.

Corré (2005) compiles a substantial number of texts featuring Lingua Franca, offering translations from the original European languages and a selection of articles analysing and contextualising Lingua Franca. He also provides a glossary, largely derived from the *Dictionnaire* (1830) and Schuchardt’s seminal (1909) work.

Schuchardt (1909) offers the earliest detailed historical and linguistic analysis of Lingua Franca. More recent in-depth texts include Dakhlija’s (2008) rather subjective book, Minervini’s (1996) comprehensive article that focuses predominantly on Italian and French sources, and Cifoletti’s (2004) forensic biography and analysis of Lingua Franca, again based almost exclusively on Romance language sources.

The key source authors are inevitably Haedo (1612), Broughton (1839) Pananti (1841) and Frank (1850).

Joanna Nolan

Abbreviations

HSA Hugo Schuchardt Archiv

TNA The National Archive

References

- Anonymous. 1830. *Dictionnaire de la langue franque ou petit mauresque, suivi de quelques dialogues familiers, et d'un vocabulaire de mots arabes les plus usuels; à l'usage des français en afrique*. Marseille: Feissat et Demonchy.
- Broughton, Elizabeth. 1839. *Six years' residence in algers*. London: Saunders & Otley.
- Camus Bergareche, Bruno. 1993. El estudio de la lingua franca: cuestiones pendientes. *Revue de linguistique romane* 57(227–228). 433–454.
- Cifoletti, Guido. 2004. *La lingua franca barbaresca*. Rome: Il Calamo.
- Clissold, Stephen. 1977. *The barbary slaves*. London: Edek.
- Corré, Alan D. 2005. *A glossary of lingua franca – fifth edition*. <https://minds.wisconsin.edu/bitstream/item/3920/go.html>.
- D'Aranda, Emanuel. 1662. *Relation de la captivité et libération de sieur emanuel d'aranda, jadis esclave à alger*. Brussels: Jean Mommart.
- D'Arvieux, Laurent. 1735. *Mémoires du chevalier d'arvieux*. Paris: Jean-Charles-Baptiste Desespine.
- Dakhli, Jocelyne. 2008. *Lingua franca*. Arles: Actes Sud.
- Dan, Pierre. 1637. *Histoire de la barbarie, et de ses corsaires divisée en six livres*. Paris: P. Rocolet.
- Davis, Robert C. 2004. *Christian slaves, muslim masters: white slavery in the mediterranean, the barbary coast and italy, 1500–1800*. London: Palgrave Macmillan.
- Dendien, Jacques. 1994. *Trésor de la langue française informatisé*. <http://stella.atilf.fr/>.
- Duclos, Jeanne. 1992. *Dictionnaire du français d'algérie: français colonial, pataouète, français des pieds-noirs*. Paris: Editions Bonneton.
- Frank, Louis. 1850. *Algérie, états tripolitains, Tunis*. Paris: Firmin Dido.
- Garcès, Maria Antonia. 2011. *An early modern dialogue with Islam: Antonio de Sosa's Topography of Algiers (1612)*. Trans. by Diana de Armas Wilson. Indiana: University of Notre Dame.

- Gardner-Chloros, Penelope. 2009. *Code-switching*. Cambridge: Cambridge University Press.
- Giancarli, Gigio Artemio. 1545. *La zingana*. Mantova: V. Ruffinello.
- Gilson, Erika H. 1987. *The turkish grammar of thomas vaughan: ottoman turkish at the end of the XVIIIth century*. Wiesbaden: Harrassowitz.
- Grandchamp, Pierre. 1920. *La France en tunisie à la fin du XVIe siècle (1582–1600)*. Tunis: Société Anonyme de l’Imprimerie Rapide.
- Grimon, G. 1890. Farmacopea e lingua franca del dugento. *Archivio Glottologico Italiano* 12. 181–186.
- Haedo, Diego de. 1612. *Topographia e historia general de argel*. Valladolid.
- Hall, Robert A. Jr. 1966. *Pidgin and creole languages*. Ithaca, New York: Cornell University Press.
- Hopkins, J. F. P. 1982. *Letters from barbary 1576–1744: arabic documents in the public record office*. Oxford: Oxford University Press.
- Kahane, Henry & Renée Kahane. 1976. Lingua franca: the story of a term. *Romance Philology* 30. 25–41.
- Minervini, Laura. 1996. La lingua franca mediterranea: plurilinguismo, mistilinguismo, pidginizzazione sulle coste del mediterraneo tra tardo medioevo e prima età moderna. *Medioevo Romanzo* 30. 231–301.
- Molière, Jean-Baptiste Poquelin. 1667. *Le sicilien*. Paris: Jean Ribov.
- Molière, Jean-Baptiste Poquelin. 1798. *Le bourgeois gentilhomme*. Marseille: Jean Mossy.
- Mori, Laura. 2016. Plurilinguismo, interferenza e marche acquisizionali in “italiano di contatto” nella comunicazione transculturale del mediterraneo moderno. In Cristina Muru & Margherita Di Salvo (eds.), *Dragomanni, sovrani e mercanti: pratiche linguistiche nelle relazioni politiche e commerciale del mediterraneo moderno*, 23–72. Pisa: Edizione ETS.
- Nolan, Joanna. 2018. *Fact and fiction: a re-evaluation of lingua franca*. London: SOAS University of London doctoral dissertation.
- Pananti, Filippo. 1841. *Avventure e osservazioni sopra le coste di barberia*. Milan: Leonardo Ciardetti.
- Pellegrini, Giovanni Battista. 1972. *Gli arabismi nelle lingue neolatine con speciale riguardo all’Italia*. Brescia: Padeia.
- Pennell, Richard. 1982. Tripoli in the mid eighteenth century: A guidebook to the city in 1767. *Revue d’histoire maghrébine* 25–26. 91–121.
- Pihan, Antoine Paulin. 1847. *Glossaire des mots français de l’arabe, du persan et du turc*. Paris: Benjamin Duprat.

Joanna Nolan

- Plantet, Eugène. 1889. *Correspondance des deys d'alger avec la cour de France, 1579–1833*. Paris: F. Alcan.
- Poiret, Jean Louis Marie. 1802. *Voyage en barbarie ou lettres écrites de l'ancienne numidie*. Paris: Callixte Volland.
- Rehbinder, Johan von. 1800. *Nachrichten und Bemerkungen über den algerischen Staat 1798–1800*. Altona: Hammerich.
- Rossetti, Roberto. 1999. An introduction to lingua franca. *Englishes: letterature inglesi contemporanee* 8. 42–62.
- Savary de Brèves, François. 1628. *Relation des voyages de monsieur de brèves, tant en grèce, terre sainte et aegypte, qu'aux royaumes de Tunis et arger*. Paris: Nicolas Gasse.
- Schuchardt, Hugo. 1909. *Pidgin and Creole languages*. Cambridge: Cambridge University Press. Trans. by Glen G. Gilbert, 1980.
- Selbach, Rachel. 2008. The superstrate is not always the lexifier: lingua franca in the barbary coast 1530–1830. In Suzanne Michaelis (ed.), *Roots of creole structures: weighing the contribution of substrates and superstrates*, 29–58. Amsterdam: John Benjamins.
- Tinniswood, Adrian. 2010. *Pirates of barbary: corsairs, conquests and captivity in the 17th century*. London: Vintage Books.
- Trivellato, Francesca. 2009. *The familiarity of strangers: the sephardic diaspora, livorno, and cross-cultural trade in the early modern period*. New Haven: Yale University Press.
- Tully, Miss. 1819. *Ten years' residence at the court of tripoli: letters written during a ten years' residence at the court of tripoli 1783–1995, vol. 1*. London: H. Colburn.
- Ursini, Flavia. 2011. *Dialetti veneti*. In Raffaele Simone (ed.), *Enciclopedia italiana*. Rome: Treccani.
- Weiss, Gillian. 2011. *Captives and corsairs: france and slavery in the early modern mediterranean*. Stanford: University of California Press.

Part III

Domains of contact-induced change across Arabic varieties

Chapter 15

Dialect contact and phonological change

William M. Cotter

University of Arizona

This chapter examines phonological and phonetic changes that have been documented and analyzed in spoken Arabic varieties, occurring as a result of dialect contact. The factors contributing to dialect contact in Arabic-speaking communities vary, from economic migration which has encouraged individuals to move into new dialect areas seeking work, to migration that stems from political violence and upheaval. These diverse factors have contributed to the large-scale migration of Arabic speakers to other parts of the Arabic speaking world. As a result, dialect contact is rampant, and decades of Arabic sociolinguistic research have shown that the phonological and phonetic outcomes of these contact situations have been quite profound.

1 Introduction

In this chapter, I discuss research that has examined the outcomes of Arabic dialect contact and the influence of contact on phonological change in spoken Arabic varieties. This chapter also discusses the interface between phonology and phonetics, and the effect of contact on these areas of the linguistic system. Given space constraints, I discuss only a portion of the published work in these areas, giving some priority to recent doctoral dissertations that have contributed to this body of research. Further, I exclude work that has investigated the effects of contact on the morphology and syntax of Arabic (e.g. [Al-Wer et al. 2015](#); [Gaftar & Horesh 2015](#); Leddy-Cecere, this volume; Lucas, this volume; Manfredi, this volume).

Although Arabic sociolinguistics is an increasingly robust area of linguistic research, limiting my discussion to cases of contact-induced phonological and phonetic change is perhaps unsurprising, given the scholarly history of dialect-contact research and its place within sociolinguistics. Sociolinguistics has made



William M. Cotter

great progress towards the goal of analyzing the full scope of variation in languages around the world. However, historically, and to some extent still today, examinations of variation and change in the realms of phonology and phonetics have been the meat and potatoes of sociolinguistic work. I would argue that this is true of Arabic sociolinguistic work as well.

From Labov's (1963) early work on Martha's Vineyard, phonetics and phonology have been at the heart of analyses of dialect contact. As a result, much of what we know about Arabic dialect contact has stemmed from earlier foundational research on dialect contact in the English-speaking world. Within this work on English, research by Milroy (1987), Trudgill (1986; 2004), Britain (2002), and Britain & Trudgill (2009), among many others, has shown how dialect contact often plays out, and how that contact influences language variation and change.

However, research on Arabic has moved beyond simply testing the hypotheses put forward by scholars of English dialect contact, playing its own role in refining sociolinguistic theory. Notably, Arabic sociolinguistics has refined our understanding of diglossia (Ferguson 1959). Ibrahim (1986) and Haeri (2000) have reoriented our understanding of Arabic diglossia from Ferguson's High-Low dichotomy to one that draws on locally meaningful understandings of linguistic prestige. In doing so, this work has moved our discussion away from analyzing Arabic through the lens of "standard" or "nonstandard" varieties or variants, setting the stage for decades of research that has examined contact-induced change in Arabic varieties.

1.1 The potential limitations of borrowing and imposition for Arabic dialect contact research

Before moving on to a discussion of a number of specific cases of Arabic dialect contact, I briefly address the potential limitations of Van Coetsem's (1988; 2000) framework for discussions of dialect contact, as opposed to language contact. After discussing Van Coetsem's approach, I shift my focus to discuss Arabic dialect contact through a theoretical lens that has proven productive in earlier sociolinguistic work (Trudgill 1986; 2004).

In analyzing phonological change as a result of dialect contact, Van Coetsem's framework presents a number of possible challenges. One specific issue is that in many cases, a clear distinction between the borrowing or imposition of linguistic forms is challenging to establish in cases of Arabic dialect contact. Scholars may encounter challenges in attempting to assert the agentivity of the recipient language in making a case for the borrowing of, for example, aspects of a

15 *Dialect contact and phonological change*

dialect's phonology into the phonology of another dialect. Asserting the agency of the source language in making the case for imposition is similarly challenging. These challenges stem from the cognitive orientation of Van Coetsem's framework, which, as Lucas (2015: 521) notes, is not based on social realities or variation in the power and prestige that a given dialect or language may hold.

The approach that many scholars within sociolinguistics and allied fields like linguistic anthropology have taken is, in contrast, inherently social. We concern ourselves with the social life of language, and although we do not discredit cognitive approaches to language acquisition and use, in much of the work on dialect contact, we have foregrounded social factors in our analyses of language change. However, it is worth noting that within sociolinguistic research on second dialect acquisition, researchers have highlighted the role of social factors, as well as the constraints placed on acquisition by the linguistic system (e.g. Nycz 2013; 2016).

With the above discussion in mind, I argue that Van Coetsem's framework is less readily applicable to the cases that I describe in this chapter. Instead, I suggest that outcomes of Arabic dialect contact are better analyzed through the framework advocated for within sociolinguistics. It is to that framework that I now turn.

As Trudgill (2004) notes in discussing new dialect formation, dialect contact often progresses in stages. One of the earliest stages in this process is *LEVELING* (Trudgill 2004: 83), which results in the reduction of forms from a given dialect. These forms may be, but do not have to be, socially marked, e.g. affrication of /k/ to [č] in certain Arabic dialects. Most importantly for Trudgill, during leveling certain variants of a given feature will supplant others (Trudgill 2004: 85). As a result, forms that are socially marked may be leveled out, while unmarked forms may survive even if they were not a majority variant. In those cases where socially marked forms are present, they are often reduced across generations. Trudgill also describes processes of interdialect development, where forms arise out of the interaction between dialects, such as reallocation, where surviving forms are reallocated in some way, and focusing, whereby a new variety born out of contact begins to stabilize (Trudgill 2004).

What I feel that this framework offers in discussions of Arabic dialect contact is an acknowledgement of the social issues that may influence linguistic change, especially in situations of contact. In the remainder of this chapter I discuss cases of contact-induced change in Arabic varieties. In doing so, I draw on sociolinguistic understandings of how contact-induced changes take hold and progress.

William M. Cotter

2 Contact-induced changes in the phonology of Arabic dialects

When discussing Arabic dialect contact, a brief discussion of the typology of Arabic is useful, as it provides a shared lexicon for discussing the outcomes of contact. Cadora (1992) offers an ecolinguistic taxonomy of Arabic, describing a continuum of Arabic varieties containing linguistic features ranging from what he describes as bedouin in provenance, to those that can be considered sedentary. In presenting a related contrast, Cadora describes features that situate dialects as being urban versus those that are rural.

However, what Cadora offers is not a hard and fast classification of Arabic varieties. Instead, his typology highlights linguistic features that typically group together within dialect types, providing a way to conceptualize the similarities and differences across these varieties. Importantly for this chapter, sites of contact between Arabic dialects are also often sites of contact between *types* of dialects as well.

In my own work on Palestinian Arabic this has been the case, with Gaza City offering one example of contact between different Palestinian Arabic dialects. Today, the dialect of Gaza City has both bedouin and urban sedentary features, dialect types that likely came into contact in Gaza as a result of Palestinian refugee migration (see de Jong's 2000 discussion of Gaza City). This contact is undeniable given Gaza's current demographic reality, which suggests that its population is roughly 70% refugee.¹ It is also unsurprising given that Gaza has long been a site of contact. This history of contact has resulted in a city dialect that looks different than other urban Palestinian varieties spoken in major cities like Jerusalem or Nablus.

The above example serves as a way to frame the linguistic discussion of contact-induced phonological change provided below. I begin by covering documented consonantal changes that have grown out of contact, before moving on to vocalic changes and the need for additional research in this area as studies of Arabic contact move forward.

¹This figure has been reported by the United Nations Relief Works Agency for Palestinian Refugees but only reflects refugees that have actually registered with the U.N. Other estimates place the percentage of Gaza's population that are refugees as closer to 80%. <https://www.unrwa.org/where-we-work/gaza-strip>.

2.1 Consonantal changes

One of the most widely discussed linguistic features within work on Arabic dialect contact has been the variable realization of the voiceless uvular stop /q/. Motivation for the scholarly interest in /q/ likely stems from a number of factors. First, the phoneme has a wide range of dialectal variation, with dialectal realizations including a true voiceless uvular [q], as well as [k, g, ʔ] and an additional [k] variant articulated between a velar and uvular (Shahin 2011). Second, interest in /q/ is also likely due to the high social salience of its variation in many Arabic-speaking communities (see Hachimi 2012; Cotter & Horesh 2015).

The result is that /q/ has been one of the most heavily studied features in Arabic sociolinguistics. Variation and change in /q/ has been discussed in a number of different communities throughout the Arabic speaking world, including: Palestine (Abd El-Jawad 1987; Al-Shareef 2002; Cotter & Horesh 2015; Cotter 2016); Egypt (Haeri 1997); Iraq (Blanc 1964; Abu-Haidar 1991);² Jordan (Abd El-Jawad 1981; Al-Wer 2007; Al-Wer & Herin 2011); Morocco (Hachimi 2007; 2012); and Bahrain (Holes 1987), among others.

What these cases suggest are robust processes of linguistic change in the realization of /q/ coming as a result of factors such as migration and dialect contact. While the social patterning of these changes (e.g. stratified along age, gender, or sectarian lines) have been as diverse as the communities in which /q/ has been analyzed, across these contexts we see regular patterns of change in /q/ over time.

Taking the case investigated by Cotter (2016) as an example, we can see how patterns of change in /q/ may progress over time. In the speech of Jaffa Palestinian refugees in the Gaza Strip, Cotter (2016) showed that across three generations Jaffa refugees in Gaza showed progressively lower use of their traditional [ʔ] realization of /q/, instead beginning to favor the voiced velar [g] variant that is common in Gaza City Arabic. Within the oldest generation of this community, Jaffa refugees showed near categorical retention of the glottal variant, and little rudimentary leveling. However, the second generation of Jaffa refugees showed substantial variability between [ʔ] and [g], while in the third generation in the study, speakers showed higher rates of usage of the [g] variant that is native to the Gaza City dialect.

However, as Cotter & Horesh (2015) discuss, variability in /q/ is often situated within broader identity projects that speakers and communities have under-

²Both Abu-Haidar and Blanc's analyses are dialectological and descriptive in scope, however both ultimately discuss what appear to be processes of change taking place for /q/ within what Blanc termed "communal" (i.e. religio-sectarian) varieties of Arabic in Baghdad.

William M. Cotter

taken. It is important then that analyses of Arabic dialect contact also consider the broader ethnographic context in which this contact takes place.

Another area of interest for researchers examining dialect contact has been the interdental fricatives /θ, ð, ð/. Across Arabic varieties, these phonemes quite often vary between realizations as true interdental fricatives [θ, ð, ð] and their stop counterparts [t, d, ɗ] (Al-Wer 1997; 2003; 2011). In addition to descriptive work that has documented the realization of the interdentals across Arabic varieties, they have also been examined as sociolinguistic variables in cases of dialect contact.

For example, Holes (1987; 1995) investigated sociolinguistic variation in the realization of /θ, ð, ð/ roughly split along sectarian lines in the speech of Arab and Baḥārna speakers in Bahrain. In Bahrain, in the dialect of Sunni Arabs these phonemes are traditionally pronounced as [θ, ð, ð], whereas in the dialect of Shi'i speakers they are pronounced as [f, d, ɗ]. Holes (1995: 275) details that in the speech of young literate speakers in Manama, intercommunal dialect realizations of the interdentals have emerged that are generally centered on the Sunni Arab realizations of these phonemes. More recently, Al-Essa (2008) examined the interdentals in the speech of Najdi Arabic speakers living in Jeddah, Saudi Arabia, an Urban Hijazi Arabic dialect area. Although Najdi Arabic typically retains the interdental realization of these phonemes, Al-Essa concluded that degree of contact with Urban Hijazi speakers was a significant factor influencing whether Najdi speakers adopted the stop realizations common in Urban Hijazi Arabic.

Additionally, Alghamdi (2014) investigated the interdentals through the lens of migration and contact in the Saudi Arabian city of Mecca. Alghamdi describes what may be the beginning of a change from the traditional interdental realization of these phonemes in the direction of their stop counterparts. As Alghamdi (2014: 112) notes, if it is the case that an incipient change in the interdentals exists in Mecca, the results of her study suggest that female speakers may be leading this change. This finding supports earlier sociolinguistic work, which has highlighted that female speakers are often at the vanguard of linguistic change.

Change in the interdentals has also been examined as part of new dialect formation in the Jordanian capital of Amman. As Al-Wer (2007) describes (see also Al-Wer, this volume), Ammani Arabic has grown out of contact between speakers of two different dialect types: urban Palestinian and traditional Jordanian varieties, which differ in their realizations of the interdentals. Urban Palestinian dialects typically favor non-interdental realizations [t, d, ɗ], while, in contrast, traditional Jordanian dialects retain the interdentals [θ, ð, ð]. Al-Wer describes the case of the interdentals in Amman as a process of focusing (Trudgill

15 *Dialect contact and phonological change*

2004) that has arisen out of contact. In Trudgill's terms (drawing on Le Page & Tabouret-Keller 1985), focusing is one part of the process of new-dialect formation, whereby features of input dialects are leveled and stability emerges, resulting in new shared linguistic norms. Al-Wer describes that, in Amman, focusing of the interdental fricatives in the direction of their stop counterparts [t, d, ð] has taken place (Al-Wer 2007: 66). In addition, Al-Wer notes that, as a result of contact, Ammani Arabic has also focused towards the common Palestinian [ʒ] realization of /ǧ/ at the expense of the traditional Jordanian [ǧ] (Al-Wer 2007: 66).

In addition to the work by Al-Essa (2008) and Alghamdi (2014) discussed above, a more recent case of contact-induced change in Saudi Arabia has been identified: the voiced lateral fricative [ɮʕ] realization of <ض>. Al-Wer & Al-Qahtani (2016) investigate /ɮʕ/ as a variable in the dialect of Tihāmat Qaḥtān. What this work shows is that in the Tihāmat Qaḥtān variety, the lateral [ɮʕ] represents a conservative, traditional variant of the phoneme, whereas the voiced interdental fricative [ð] represents the innovative variant. As a result of dialect contact, Al-Wer & Al-Qahtani (2016) describe an intergenerational process of change towards the voiced emphatic interdental [ð], with use of the historic [ɮʕ] variant receding over time.

Another area of interest in dialect contact research has been affrication. As descriptive work has shown, affrication of certain phonemes, notably /k/ in the direction of [č], is common in Arabic dialects. As an example of this process, Shahin (2011) notes that in rural varieties of Palestinian Arabic, /k/ palatalizes to become an affricate [č] (e.g. *čīfak* 'how are you (SG.M)?' < *kīfak*). While typologically this affrication is common, processes of affrication or de-affrication have also been noted as the outcome of contact.

Al-Essa (2008) investigated affrication of /k/ and /g/ in the speech of Najdi Arabic migrants in Jeddah, and found that the affrication that is a common feature of this variety had been almost completely undone in this migrant community. Examining this change in light of dialect contact, Al-Essa concludes that this deaffrication represents the leveling out of marked regional dialect forms as a result of contact (Trudgill 1986; Kerswill & Williams 2000). More recently, Al-Wer et al. (2015) note that the conditional, root-based distribution of the affricate [č] for /k/ in the Sult variety of Arabic in Jordan, which, although it has receded (Al-Wer 1991), now interacts with other innovative features in Sult that show potential stratification along religious lines.

Elsewhere in Jordan, notably in Amman, Al-Wer (2007) describes the leveling of the affricate [č] across generations. The city dialect that has emerged in Amman, which has Sulti Arabic as one of its input varieties, underwent rudimentary

William M. Cotter

leveling (Trudgill 2004) within the first generation. This leveling resulted in the loss of this affricate variant of /k/ in the speech of Sulti migrants. In this case, Al-Wer describes the deaffrication of [č] as stemming from its status as a marked feature of Horani Arabic varieties like that of Sult. This marked status makes it a primary candidate for the kinds of leveling that sociolinguists have identified in other cases of contact.

2.2 Vocalic changes

In general, the Arabic vocalic system remains understudied within research on Arabic varieties. However, multiple cases of change linked to dialect contact have been identified. One of the most well studied cases of contact-induced vocalic change in Arabic is perhaps better thought of as a morphophonological change: the Arabic feminine gender marker. The feminine gender marker is a word final vocalic morpheme that is realized variably across Arabic varieties. The realization of this vowel varies from an unraised [a] to [æ, ɛ, e], or even as high as [i] (e.g. Al-Wer 2007; Naïm 2011; Shahin 2011; Woidich 2011).

Even within one region, the full range of variation in this morpheme can be seen. Taking the Levant as an example, the Lebanese capital of Beirut is known for raising this vowel to [e] or even [i] (Naïm 2011). The Syrian capital of Damascus is known to raise to [e] (Lentin 2011). Urban Palestinian (Rosenhouse 2011; Shahin 2011) is also often described as raising to [e], while Amman (Al-Wer 2007) raises this vowel to [ɛ]. These city varieties can be contrasted with, for instance, the variety of Cairo (Woidich 2011), which does not raise this vowel, leaving it as [a].

This morpheme is particularly interesting within a discussion of dialect contact because raising of this vowel is phonologically conditioned. The phonological factors that constrain raising vary across dialects, with urban Levantine Arabic (e.g. Syria, Palestine, Lebanon) providing one example of these factors. In urban Levantine, the following rules constrain the raising of this vowel (Grotzfeld 1980: 181; Levin 1994: 44-45; Al-Wer 2007: 68):

1. The default realization of the vowel is raised: [e]
2. The vowel is unraised, realized as [a], when:
 - a) it occurs after back consonants (i.e. pharyngeal, glottal, emphatic/pharyngealized, post-velar): /ħ, ʕ, ʔ, h, ʁ, d/ð, ɰ, ʁ, q/;
 - b) it occurs after /r/, but only when preceding /r/ there is no high front vowel. In cases where a high front vowel does precede /r/, raising is

15 *Dialect contact and phonological change*

allowed, e.g. [kbi:re] ‘big (F)’.

Below I provide two specific documented examples of contact and change in the feminine gender marker. First, [Cotter & Horesh \(2015\)](#) investigated change in the feminine gender marker in the speech of refugees originally from the Palestinian city of Jaffa who now live as refugees in the Gaza Strip. This sample included both speakers who were expelled from Jaffa after the creation of the state of Israel in 1948 and their descendants. Their traditional urban Palestinian dialect ([Horesh 2000](#); [Shahin 2011](#)) is one that raises the feminine gender marker to [e], subject to the phonological conditioning mentioned above. In contrast, based on the available dialectological information the dialect of Gaza City does not raise this vowel ([Bergsträßer 1915](#)).

[Cotter & Horesh \(2015\)](#) highlight a process of contact-induced change that has taken place in this community. Across generations, the realization of this vowel appears to be lowering and backing, moving from [e] in the direction of [a]. The result is that younger Jaffa refugee speakers realize the vowel closer to the [a] common in Gaza City. This type of change is perhaps unsurprising in a city like Gaza, given that the population of Gaza is overwhelmingly comprised of refugees, including large communities who are of [a] dialect types for this feature. This diversity and the high numbers of refugees in Gaza means that the city, and the territory generally, is a site where many dialects of Palestinian Arabic are in intimate contact. What remains to be determined is whether or not new linguistic norms are emerging in the dialect of Gaza City more generally as a result of this contact.

One other case, which is discussed in more detail by Al-Wer (this volume), provides a succinct example of the intersection between phonetics, phonology, and Arabic dialect contact. In discussing the formation of Ammani Arabic, [Al-Wer \(2007\)](#) notes the centrality of vocalic change to the formation of the dialect. The feminine gender marker represents one feature that has helped to define the variety of Amman.

As [Al-Wer \(2007\)](#) describes, through contact between Palestinian and Jordanian Arabic dialects in Amman, the realization of the feminine gender marker has focused on [ɛ], the indigenous Jordanian realization (as in e.g. the Horani dialect of Sult, see [Herin 2014](#)). However, although Ammani Arabic has focused on the Jordanian phonetic realization of this vowel, it has retained urban Palestinian phonology, which blocks raising in the environment of back consonants as defined above ([Al-Wer 2007](#): 69). This is less restrictive than in Horani Arabic, where raising is also blocked in the environment of velar consonants such as /k/ and the labiovelar /w/ ([Al-Wer et al. 2015](#): 77).

William M. Cotter

Finally, I mention one other case of vocalic change that has been documented as an outcome of dialect contact: the diphthongs [ay] and [aw]. [Alghamdi \(2014\)](#) investigated monophthongization of the traditional Arabic diphthongs /ay/ and /aw/ in the speech of Ghamdi migrants in Mecca. Alghamdi found that the diphthongs common in the dialect spoken by this migrant community were monophthongizing, reflecting a change towards the norms of Mecca Arabic, which lacks diphthongs. Alghamdi's analysis of the diphthongs provides an example of dialect leveling borne out of contact, noting two additional aspects of this variable in the speech of this migrant community: i) Alghamdi describes the high degree of social salience that the diphthongs have in this community and their possible stigmatization in Mecca, and ii) that retention of the diphthongs is uncommon in Saudi Arabic, making the Ghamdi realization a minority realization in Saudi Arabia generally. These two facts create an environment conducive to change.

3 Conclusion

In examining Arabic dialect contact, a growing body of research highlights that the phonology and phonetics of Arabic represent rich sites for linguistic change. As the examples that I have provided throughout this chapter, and those discussed elsewhere throughout this volume suggest, we can identify a number of cases where dialect contact has influenced the directionality and extent of change in Arabic dialects. With the findings of this selection of work in mind, a number of areas remain open for future investigation.

Perhaps the most pressing of these is the reality that, although I have highlighted work here that investigates vocalic change, the vocalic system of Arabic varieties remains drastically understudied. Although phonetic research on the vocalic system of Arabic varieties continues to grow (see e.g. [Hassan & Heselwood 2011](#); [Khattab & Al-Tamimi 2014](#); [Al-Tamimi & Khattab 2015](#)), we know little about sociophonetic changes that may take place in cases of contact like those discussed in this chapter. Given the scope of dialect contact in the Arabic-speaking world, much of which has come as a result of mass migration throughout the region, investigating the potential for processes such as vocalic chain-shifting ([Al-Wer 2007](#)) represents an important next step for research on language variation and change in Arabic. I would argue that more robust investigation of vocalic change in Arabic dialects represents a pressing area of concern for Arabic sociolinguistics.

In addition, examples like the feminine gender marker in Amman ([Al-Wer 2007](#)) open the door for future work that investigates the potential for blending

15 *Dialect contact and phonological change*

of the phonetics and phonology of different Arabic varieties as a result of contact. Although Amman is a somewhat different case, given that it represents an example of new dialect formation, a close examination of phonetics and phonology together in contact situations will provide us with an opportunity to examine how dialect focusing and leveling takes place, and how the linguistic systems of multiple different Arabic varieties interact and regularize through contact.

Additional research that looks more closely at change in the vocalic system of Arabic dialects will go a long way towards enriching the depth of Arabic sociolinguistic research. This is especially true of work that examines cases of dialect contact. However, beyond sociolinguistics, a closer examination of the vocalic system will contribute to the description and documentation of Arabic dialects, which will further enrich linguistic research that investigates the varieties of Arabic spoken around the world.

Further reading

As I have discussed throughout this chapter, Al-Wer's (2007) work details a number of aspects of how the dialect of Jordan's capital, Amman, emerged as a result of contact between Jordanian and Palestinian dialects. It offers a clear picture of how contact has played out in Amman with respect to a number of linguistic features, and how the city's dialect emerged over successive generations.

Cotter & Horesh's (2015) article examines contact in the Gaza Strip, one of the more understudied areas within Arabic linguistic research. The article draws on sociolinguistic fieldwork conducted in Gaza City, as well Jaffa and the West Bank, to analyze change in three specific features of Palestinian Arabic.

Holes' (1987) work on Bahrain provided one of the early accounts of sociolinguistic variation and change in Arabic. In addition, this work provides a clear example of variation in Arabic that has been stratified on sectarian, or religious, lines.

Acknowledgements

I would like to thank Christopher Lucas and the anonymous reviewer for their comments on this chapter. In addition, I would like to thank Uri Horesh and Enam Al-Wer for their feedback, as well as the attendees of the Arabic and Contact-Induced Change workshop at the 23rd International Conference on Historical Linguistics for their feedback on my own research as it appears in this chapter.

William M. Cotter

References

- Abd El-Jawad, Hassan. 1981. *Lexical and phonological variation in spoken Arabic in amman*. Philadelphia: University of Pennsylvania doctoral dissertation.
- Abd El-Jawad, Hassan. 1987. Cross-dialectal variation in arabic: competing prestigious forms. *Language in Society* 10. 359–67.
- Abu-Haidar, Farida. 1991. *Christian Arabic of Baghdad*. Wiesbaden: Harrassowitz.
- Al-Essa, Aziza. 2008. *Najdi speakers in Hijaz: A sociolinguistic investigation of dialect contact in jeddah*. Colchester: University of Essex doctoral dissertation.
- Al-Shareef, Jamal. 2002. *Language change and variation in palestine: a case study of jabalia refugee camp*. Leeds: University of Leeds doctoral dissertation.
- Al-Tamimi, Jalal & Ghada Khattab. 2015. Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants. *Journal of the Acoustical Society of America* 138(1). 344–60.
- Al-Wer, Enam. 1991. *Phonological variation in the speech of women from three urban areas in jordan*. Colchester: University of Essex doctoral dissertation.
- Al-Wer, Enam. 1997. Arabic between reality and ideology. *International Journal of Applied Linguistics* 7(2). 251–65.
- Al-Wer, Enam. 2003. Variability reproduced: A variationist view of the [ḍ]/[d] opposition in modern Arabic dialects. In Kees Versteegh, Martine Haak & Rudolf De Jong (eds.), *Approaches to Arabic dialects*, 21–31. Leiden: Brill.
- Al-Wer, Enam. 2007. *Arabic in the city: Issues in dialect contact and language variation*. In Catherine Miller, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.), 55–76. London: Routledge.
- Al-Wer, Enam. 2011. Phonological merger. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Al-Wer, Enam & Khairiah Al-Qahtani. 2016. Lateral fricative ḍād in Tihāmat Qahtān: A quantitative sociolinguistic investigation. In Stuart Davis & Usama Soltan (eds.), *Perspectives on arabic linguistics XXVII: papers from the Annual Symposium on Arabic Linguistics, Bloomington, Indiana, 2013*, 151–69. Amsterdam: John Benjamins.
- Al-Wer, Enam & Bruno Herin. 2011. The lifecycle of *Qaf* in Jordan. *Langage et société* 138. 59–76.
- Al-Wer, Enam, Uri Horesh, Bruno Herin & Maria Fanis. 2015. How arabic regional features become sectarian features: jordan as a case study. *Zeitschrift für Arabische Linguistik* 62. 68–87.

15 *Dialect contact and phonological change*

- Alghamdi, Najla Manie. 2014. *A sociolinguistic study of dialect contact in arabia: ghamdi immigrants in mecca*. Colchester: University of Essex doctoral dissertation.
- Bergsträßer, Gotthelf. 1915. *Sprachatlas von Syrien und Palästina*. Leipzig: Hinrichs.
- Blanc, Haim. 1964. *Communal dialects in baghdad*. Cambridge, MA: Harvard University Press.
- Britain, David. 2002. Diffusion, levelling, simplification and reallocation in past tense *BE* in the English Fens. *Journal of Sociolinguistics* 6(1). 16–43.
- Britain, David & Peter Trudgill. 2009. New dialect formation and contact-induced reallocation: three case studies from the English Fens. *International Journal of English Studies* 5(1). 183–209.
- Cadora, F. J. 1992. *Bedouin, village and urban Arabic: an ecolinguistic study*. Leiden: Brill.
- Cotter, William M. 2016. (q) as a sociolinguistic variable in the Arabic of Gaza City. In Youssef A. Haddad & Eric Potsdam (eds.), *Perspectives on arabic linguistics XXVIII: papers from the annual symposium on arabic linguistics, gainesville, florida, 2014*, 229–246. Amsterdam: John Benjamins.
- Cotter, William M. & Uri Horesh. 2015. Social integration and dialect divergence in coastal palestine. *Journal of Sociolinguistics* 19(4). 460–83.
- De Jong, Rudolf. 2000. *A grammar of the bedouin dialects of the northern sinai littoral: bridging the linguistic gap between the eastern and western arab world*. Leiden: Brill.
- Ferguson, Charles A. 1959. Diglossia. *Word* 15. 325–40.
- Gafer, Roey & Uri Horesh. 2015. When the construction is *axla*, everything is *axla*: a combined lexical and structural borrowing from Arabic into Hebrew. *Journal of Jewish Languages* 3(1–2). 337–48.
- Grotzfeld, Heinz. 1980. Das syrisch-palästinensische arabisch. In Wolfdietrich Fischer & Otto Jastrow (eds.), *Handbuch der arabischen Dialekte*, 174–90. Wiesbaden: Harrassowitz.
- Hachimi, Atiqa. 2007. Becoming casablanca: Fessis in casablanca as a case study. In Catherine Miller, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.), *Arabic in the city: Issues in dialect contact and language variation*, 97–122. London: Routledge.
- Hachimi, Atiqa. 2012. The urban and the urbane: identities, language ideologies, and arabic dialects in morocco. *Language in Society* 41(3). 321–41.
- Haeri, Niloofar. 1997. *The sociolinguistic market of Cairo: Gender, class and education*. London: Routledge.

William M. Cotter

- Haeri, Niloofar. 2000. Form and ideology: arabic sociolinguistics and beyond. *Annual Review of Anthropology* 29(1). 61–87.
- Hassan, Zeki Majeed & Barry Heselwood. 2011. *Instrumental studies in Arabic phonetics*. Amsterdam: John Benjamins.
- Herin, Bruno. 2014. Salt, dialect of. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Holes, Clive. 1987. *Language variation and change in a modernising Arab state: The case of Bahrain*. London: Kegan Paul.
- Holes, Clive. 1995. Community, dialect and urbanization in the arabic-speaking Middle East. *Bulletin of the School of Oriental and African Studies* 58(2). 270–87.
- Horesh, Uri. 2000. Toward a phonemic and phonetic assessment of jaffa Arabic: Is it a typical urban palestinian dialect? In Manwel Mifsud (ed.), *Proceedings of the third international conference of AIDA: Association Internationale de Dialectologie Arabe, held in Malta, 29 March – 2 April 1998*, 15–20. Malta: AIDA.
- Ibrahim, Muhammad H. 1986. Standard and prestige language: a problem in arabic sociolinguistics. *Anthropological Linguistics* 28(1). 115–26.
- Kerswill, Paul & Ann Williams. 2000. Creating a New Town Koine: Children and language change in Milton Keynes. *Language in Society* 29(1). 65–115.
- Khattab, Ghada & Jalal Al-Tamimi. 2014. Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure. *Laboratory Phonology* 5(2). 231–269.
- Labov, William. 1963. The social motivation of a sound change. *Word* 19(3). 273–309.
- Le Page, R. B. & Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- Lentin, Jérôme. 2011. Damascus Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Levin, Aryeh. 1994. *A grammar of the Arabic dialect of Jerusalem*. Jerusalem: Magnes Press.
- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Milroy, Lesley. 1987. *Language and social networks*. Oxford: Wiley.
- Naïm, Samia. 2011. Beirut arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Nycz, Jennifer. 2013. Changing words or changing rules? second dialect acquisition and phonological representation. *Journal of Pragmatics* 52. 49–62.

15 *Dialect contact and phonological change*

- Nycz, Jennifer. 2016. Awareness and acquisition of new dialect features. In Anna Babel (ed.), *Awareness and Control in Sociolinguistic Research*, 62–79. Cambridge: Cambridge University Press.
- Rosenhouse, Judith. 2011. Jerusalem Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Shahin, Kimary. 2011. Palestinian Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Wiley-Blackwell.
- Trudgill, Peter. 2004. *New-dialect formation: The inevitability of colonial Englishes*. Oxford: Oxford University Press.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Woidich, Manfred. 2011. Cairo arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Chapter 16

Contact-induced grammaticalization between Arabic dialects

Thomas Leddy-Cecere

Bennington College

This chapter describes the phenomenon of contact-induced grammaticalization between Arabic dialects and its significance in accounting for the development of future tense markers across modern Arabic varieties. After an introduction to theoretical aspects of general grammaticalization theory and contact-induced grammaticalization in particular, discussion shifts to the identification of specific contact-induced grammaticalization processes leading to the modern distribution of future tense-marking forms across the Arabic-speaking world. Finally, the significance of these findings to broader inquiry in Arabic dialectology and theoretical contact linguistics is considered.

1 Introduction & theory

1.1 Overview

This chapter presents evidence for the occurrence of contact-induced grammaticalization processes between dialects of Arabic over the course of the language's history. The critical role of dialect contact as a source of synchronic variation and diachronic change across Arabic varieties is well recognized, and the description of its outcomes a long-standing occupation of Arabic dialectologists (e.g. [Behnstedt & Woidich 2005](#); [Miller et al. 2007](#)). Representing a fairly recent theoretical development in the field of contact linguistics – following largely from the proposals of [Heine & Kuteva \(2003; 2005\)](#) and [Dahl \(2001\)](#) – contact-induced grammaticalization as a model has not been applied to the analysis of Arabic dialect data on a large scale. As will be seen, however, the phenomenon displays



Thomas Leddy-Cecere

significant merit as an account for the evolution and diffusion of a number of morphosyntactic features across the modern Arabic dialects.

In the following subsections, I begin with a review of the current state of research in grammaticalization theory and in contact-induced grammaticalization (CIG) specifically. I then proceed to an illustrative example of CIG in the Arabic context, demonstrating the model's power as an explanatory mechanism in interpreting the distribution of future tense markers across the modern dialects. I conclude the chapter with a brief discussion of the broader significance of CIG in the analysis of Arabic and the potential role for Arabic data in advancing general theoretical knowledge of the phenomenon at large.

1.2 Grammaticalization

Most linguists agree that it is possible to synchronically classify the majority of linguistic forms along a cline from “more lexical” to “more grammatical”, in a manner roughly consistent with the progression as conceived by [Hopper & Traugott \(2003\)](#):

CONTENT WORD > GRAMMATICAL WORD > CLITIC > INFLECTIONAL AFFIX

Historical linguists would add to this synchronic observation the diachronic reflection that it is common to observe a single etymological item advancing through the successive stages of this cline as it develops as part of a linguistic system over time. In fact, the sheer frequency of examples indicating such a trajectory of evolution has led to the identification of a cross-linguistically attested phenomenon known as grammaticalization. The recent definition provided by Hopper and Traugott is indicative of several currently referenced in the field, which – though differing in emphasis and points of detail – broadly subscribe to a similar central principle:

[Grammaticalization is] the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions ([Hopper & Traugott 2003: 18](#)).

Though useful for purposes of general definition, this largely intuitive formulation of grammaticalization and the cline which it follows must be further deconstructed if they are to be operationalized as part of a rigorous analysis. [Andersen \(2008\)](#) summarizes the issue succinctly in observing that the grammaticalization

16 *Contact-induced grammaticalization between Arabic dialects*

cline so articulated conflates numerous discrete dimensions of language change by presenting them as unified steps in a chain: the shift from lexical to grammatical word is one of semantic content, while that from word to clitic to affix involves morphosyntax and any associated loss of phonological material is best understood as phonological change. Since the early stages of grammaticalization research, more complex approaches to the description of the phenomenon have been proposed based on the concurrent evaluation of multiple parameters (e.g. [Lehmann 1985](#)). Other authors opt instead to define analogous parameters in terms of diachronic processes, thereby rendering them more directly relatable to the modes of historical linguistic analysis which underlie the bulk of investigations in grammaticalization research. The latter approach is adopted here, largely following the account proposed by [Heine \(2004\)](#).

Heine views grammaticalization as defined by the simultaneous progression of four distinct but interrelated diachronic processes: desemanticization, extension, decategorialization, and erosion. Desemanticization involves the loss of concrete lexical (“content”) meaning and a corresponding rise in abstract grammatical function associated with the use of an item in particular contexts. This often represents the first observable stage of grammaticalizing change, and, as its name suggests, primarily concerns the semantic content of the item rather than its distribution, form, or syntactic behavior. Closely coupled with desemanticization is extension, namely the novel use of the grammaticalizing item in contexts in which it was not previously employed; extension is thus defined as a change in incidence. The hand-in-hand advance of these two processes is demonstrated in the evolution of the French negative element *pas*: having shed its content semantics as a noun meaning ‘step’ and developed grammatical function as a marker of verbal negation, *pas* is extended in contemporary usage to contexts involving none of the implied motion of its lexical source ([Hansen & Visconti 2009](#): 137–138).

The third process described by Heine, that of decategorialization, consists of the changes by which a grammaticalized item comes to lose the morphosyntactic properties characteristic of its source’s original word class, such as word order freedom or agreement inflection; an example may be found in the gradual development of the English adverbial marker *-ly* from a morphosyntactically free substantive meaning ‘body, form’ to a bound derivational suffix ([Ramat 2011](#): 505). Erosion, the fourth process considered by Heine, refers to the gradual reduction and lenition of phonological form beyond what is accounted for by regular sound change, as observed in the irregular changes deriving the Jewish Babylonian Aramaic continuous aspect marker *qā ~ kā* from earlier *qāʔē ‘standing’

Thomas Leddy-Cecere

(Rubin 2005: 134).

Theories of grammaticalization have also been strongly linked to the notion of unidirectionality, the proposal that change along the above-described cline occurs only from more lexical to more grammatical and not vice versa (Lehmann 2015). Though the absolute formulation of this hypothesis has been the subject of much debate (e.g. Norde 2009), recognition of a strong unidirectional tendency remains integral to understandings of grammaticalization on both empirical and theoretical grounds (Haspelmath 1998; Heine 2004). It has been proposed that the impetus for such a tendency lies in a universal set of cognitive and communicative principles common to the human mental faculty (Claudi & Heine 1986; Bybee 2003; Lehmann 2015); these would provide an account for the pervasive occurrence of grammaticalization as a worldwide phenomenon, and may be seen to bias the results of grammatical change in the directions entailed by the four processes described above.

The concomitant advancement of these processes is discernible in one of the few cases of Arabic grammaticalization for which a reasonably complete chain of historical development is attested: that of the Egyptian Arabic future tense marker *ha-*. Documented in sixteenth- and seventeenth-century sources as *rāyih*, this item already shows evidence of desemanticization and extension, departing from the semantics of its lexical source as an active participle ‘going’ to indicate intention and imminent futurity of action, consequently allowing its extension to usage contexts devoid of actual motion: *ʔanā rāyih aʔannī ʕalēh* ‘I am going to sing about it (and proceeds to sing)’ (Davies 1981: 241). In its nineteenth-century incarnation *rāḥ* ~ *raḥ* ~ *ḥa-*, the form shows further desemanticization and extension as its value changes from an imminent to a general future and it comes to be employed in previously unacceptable circumstances, such as in the presence of a non-immediate temporal adverb: *rāḥ yīgi bukra* ‘he’ll come tomorrow’ (Elias & Elias 1981: 157; for earlier usage constraints, see Davies 1981: 241-243). These increasingly modern forms also attest decategorialization, as the once-obligatory adjectival agreement marking of the participial original becomes optional – *raḥ* (SG.M)/*raḥa* (SG.F)/*raḥīn* (PL) ~ *raḥ* (invar.) (Vollers 1895: 40) – and eventually ceases to exist altogether in the tightly bound modern clitic *ha-* ~ *ḥa-* (Abdel-Massih et al. 2009: 268). Fourthly, progressive phonetic erosion is visible throughout the item’s history, as none of the loss or lenition of phonetic material through the stages *rāyih* > *rāḥ* > *raḥ* > *ḥa-* > *ḥa-* attested above is attributable to regular sound change. Taken together, these combined processes chart the grammaticalization of lexical *rāyih* ‘going’ through its gradual development into

16 *Contact-induced grammaticalization between Arabic dialects*

the modern future tense clitic *ḥa-*.¹

Having established these understandings of grammaticalization and its component processes, we now turn to the proposal that specific grammaticalization pathways are able to be shared between interacting languages and dialects: the aforementioned CIG.

1.3 Contact-induced grammaticalization

The most elaborated theoretical model proposed for CIG is that of Heine & Kuteva (2003; 2005). This model represents the phenomenon by which “a grammaticalization process [...] is transferred from the model (M) to the replica language (R),” without corresponding transfer of any actual phonological form (2003: 539). As paraphrased and clarified by Law (2014: 215), this occurs when one language, “the ‘replica language,’ develops a feature observed in another language, the ‘model’ language, but goes through a path of universal development using resources internal to the replica language.” The result is such as that seen in the Basque innovation of an allative preposition from the noun *buru* ‘head’ and a perfect tense formed with the lexical verb *ukan* ‘have’, apparently influenced by parallel grammaticalizations in neighboring varieties of Romance (Haase 1992 *apud* Heine & Kuteva 2003: 550). Such an effect is proposed to be actuated according to the following model:

1. Speakers of language R notice that in language M there is a grammatical category M_x .
2. They develop an equivalent category, R_x , using material available in their own language (R).
3. To this end, they replicate a grammaticalization process they assume to have taken place in language M, using an analogical formula of the kind $[M_y > M_x] = [R_y > R_x]$.
4. They grammaticalize R_y to R_x . (Heine & Kuteva 2003: 539)

¹On sources referenced in the preceding paragraph: Davies (1981) is a study of colloquial elements in the seventeenth century Egyptian text *Hazz al-quḥūf fī šarḥ qaṣīd Abī Šādūf*; Vollers (1895) is a descriptive grammar of Egyptian Arabic at the close of the nineteenth century, and Elias & Elias (1981) an English–Egyptian Arabic vocabulary and phrasebook first released ca. 1899; Abdel-Massih et al. (2009) is a reference grammar of modern Egyptian Arabic.

Thomas Leddy-Cecere

This proposal for the diffusion of parallel grammaticalization trajectories across linguistic varieties is presaged by Bisang's (1998) observation of the potential synergy between grammaticalization, which he views primarily as a construction-based process, and previously observed forms of contact-induced structural convergence. The phenomenon has also been influentially described by Dahl in the form of "gram families", consisting of groups of "grams [grammaticalized items] with related functions and diachronic sources that show up in genetically and/or geographically related groups of languages, in other words, what can be assumed to be the result of one process of diffusion" (Dahl 2001: 1469). Heine & Kuteva draw heavily on Dahl's theorizations, though they diverge from him in a few critical ways. First, they are significantly more conservative than Dahl in identifying examples of the phenomenon, insisting on corroborating evidence of language contact in order to posit CIG rather than inductively inferring its occurrence given genetic relatedness or proximity. Second, they do not necessarily attempt to link multiple replications of the same grammaticalization pathway into "one process of diffusion," but instead prefer to treat them as individual instances of contact between participating languages.

Further, Heine & Kuteva's model is primarily situated in the context of contact between genetically distinct languages. Regarding the occurrence of CIG between related language varieties or dialects, Dahl sees such scenarios as generating the bulk of evidence for the phenomenon: "in the majority of all such cases [of areally diffused grams], the languages involved are more or less closely related" (2001: 1469). Heine & Kuteva are wary of such identifications. Critically, however, their reasons for being so are methodological rather than theoretical. In their analysis, they choose to rely on the principle of genetic patterning as "an empirically well-founded tool for identifying cases of contact-induced linguistic transfer" (2005: 33–34), meaning that examples of CIG between unrelated languages are often easiest to identify and defend and thus have been favored in the effort to present an unambiguous account. Regarding the broader occurrence of the phenomenon, however, they state that "genetic relationship is entirely irrelevant" (2005: 184) and that CIG may occur between related languages just as it does between unrelated ones. They remain, though, more careful than Dahl to set apart cases attributable to inheritance of any stage of the grammaticalization chain from a common ancestor, which could lead to a superficially similar result not in fact dependent on any degree of contact. Along the same lines, Law (2014) reminds us that when dealing with closely related languages the possibility of drift or typological poise precipitating parallel development rises dramatically in likelihood. Thus, the analyst must be stringent in linking proposed cases of

16 *Contact-induced grammaticalization between Arabic dialects*

CIG to cross-linguistically attested paths and parameters of grammaticalization and not to the local idiosyncrasies of a given language family or subgroup.

To Heine & Kuteva, CIG is unambiguously situated in terms of Van Coetsem's (1988; 2000) dichotomy between source language (SL) and recipient language (RL) agentivity: the four-stage model of replication presented above clearly identifies speakers of the RL as the agents of contact-induced change in this instance. This judgment has opened the proposal to major critique, as several key theorists maintain that structural pattern replication of the kind required for CIG is only possible in a scenario of SL agentivity. Ross (2007), for example, views the phenomenon as part of a broader process of bilingual calquing (cf. Manfredi, this volume), involving the subconscious imposition of the functional range of an SL item onto its RL equivalent, followed secondarily by the processes Heine & Kuteva attribute to grammaticalization but which Ross views as the natural result of increases in frequency and automatization stemming from the RL item's new functionality. Ross asserts that "one cannot reasonably argue" for Heine & Kuteva's construal of CIG as an RL-agentive direct replication of a grammaticalization process because of his conviction that the phenomenon is "largely driven by effort reducing practices of which speakers are only marginally aware" (2007: 135).

Matras (2009; 2011), however, supports Heine & Kuteva's initial characterization by arguing for RL-agentivity in his own recent accounts of CIG, which provide more attention to the role played by the individual bilingual in the phenomenon's actuation. He cites the individual's communicative imperatives and creative impulse as the primary force driving the replication of grammaticalization processes, as speakers actively borrow from constructions they control in one of their languages as a source of expressive innovation in the other, limiting this transfer solely to "pattern" out of respect for the norms of the distinct speech communities in which they operate. Matras' account thus has the benefit of aligning with the motivating forces theorized to obtain for grammaticalization processes more generally. As described by Lehmann in his consideration of grammaticalization's communicative/pragmatic dimension:

To the degree that language activity is truly creative, it is no exaggeration to say that languages change because speakers want to change them ... they do not want to express themselves the same way they did yesterday, and in particular not the way that somebody else did yesterday (Lehmann 1985: 315).

Building upon this position, it holds that in scenarios of language or dialect

Thomas Leddy-Cecere

contact innovating speakers may very well wish to express themselves the same way somebody else did yesterday if the means of expression involved are novel to a distinct speech community with which they are interacting today. This synergy with less controversial understandings of grammaticalization outside the context of language/dialect contact provides a viable counterpoint to the skepticism voiced by Ross, and strongly recommends the association of CIG with RL-agentivity.

In the case of contact between closely related varieties, this characterization of CIG may be further qualified. Under Matras' RL-agentive formulation, pattern replication occurs in the presence of constraints against the usage of an SL's forms due to speaker expectation and "language loyalty" among members of the RL community (2011: 283). In the broader language contact literature, such sociolinguistic constraints have been noted to play a role in pattern diffusion (e.g. Epps 2005), and presumably interact with speakers' judgments of interlocutors' perceived bilingual competency in favoring or disfavoring matter-based mixing or borrowing (cf. Grosjean 2001). In contexts where mutual comprehensibility is a less salient concern, the drivers of pattern vs. matter-based innovation would be expected to be almost purely sociolinguistic and pragmatic, and are perhaps most fruitfully understood through the lenses of indexicality (Silverstein 2003) and focus at the level of the speech community (Le Page & Tabouret-Keller 1985) rather than as a desire to adhere to a reified linguistic code. Such is the state of affairs most likely to obtain in the case of CIG between neighboring varieties of Arabic, to which we now turn in detail.

2 CIG in the development of Arabic future tense markers

2.1 Methods of investigation

In the following subsections, I present evidence for the role of CIG as the primary mechanism underlying the development and distribution of future tense markers in the modern Arabic dialects. The data considered is drawn from a survey of eighty-one geographic sample points spanning the contiguous Arabic-speaking world, based on a total of eighty-eight descriptive sources.² This sample was constructed as part of the broader investigation of CIG between Arabic dialects presented by Leddy-Cecere (2018), which investigates the role of CIG in the development of a number of morphosyntactic features in modern Arabic varieties,

²Details of sample composition and sources consulted, as well as the selection criteria for each, may be found in Leddy-Cecere (2018: 34–35, 43–46).

16 *Contact-induced grammaticalization between Arabic dialects*

including future tense markers, genitive exponents, and temporal adverbs meaning ‘now’. A discussion of data and findings for the first of these features is the focus of this chapter, and these shall be seen to argue strongly for the identification of CIG as a key force in shaping the evolution of modern Arabic dialects. Readers are encouraged to refer to [Leddy-Cecere \(2018\)](#) for additional examination and expansion of the points to follow.

To begin, I first examine the complete set of specific grammaticalizations of future tense markers attested by the Arabic dialect data. These have been identified via the observation of concurrent processes of desemanticization, extension, decategorialization and erosion (as described in §1.2). These individual instances are further sorted into higher level groupings by grammaticalization path: future tense markers deriving from motion verbs meaning ‘go’, for example, as against those deriving from purposive constructions, etc. Special attention is paid to those specific grammaticalization paths represented by multiple evolutions involving distinct etyma, thus identifiable as potential candidates for the products of replication through CIG. As a final step in the evaluation, the geographic incidence of forms representing such multiply attested paths is considered to assess whether their modern distribution is consistent with a historical account of diffusion via contact. This latter portion of the analysis will be presented in (§2.3), following the complete accounting of grammaticalized forms provided in (§2.2) immediately below.³

2.2 Grammaticalizations of Arabic future tense markers by grammaticalization path

2.2.1 Futures from ‘go’ (FUT < GO)

Grammaticalizations of future tense markers from forms of lexical verbs meaning ‘go’ are well represented in the Arabic data. This grammaticalization path is widely attested cross-linguistically, providing one of the major sources for the development of future tense markers worldwide ([Bybee et al. 1994](#); [Heine & Kuteva 2002](#)). Grammaticalizations of specific items observed in the cross-dialectal Arabic sample are described below.

*rāyih

Future markers representing grammaticalized forms of an active partici-

³For further discussion and justification of each stage of this heuristic, see [Leddy-Cecere \(2018: 36–43, 209–214\)](#).

Thomas Leddy-Cecere

ple *rāyih ‘going’ are found across a broad east-west swath of the Arabic-speaking world, extending from southern Iraq in the east to Algerian territory in the west. Differing degrees of grammaticalization are attested, with some forms maintaining full phonological integrity and categorial membership (e.g. Basra *rāyih*; Mahdi 1985, which displays adjectival gender/number agreement with its subject) and others showing dramatic erosion and of loss of morphosyntactic autonomy (including Cairo *ha-* ~ *ha-*, as described in §1.2). Semantically, some forms, such as Algiers *rāḥ* and Jerusalem *rāyih* ~ *rāḥ* ~ *ḥā-*, are recorded as expressing a value of immediate future or future intent (Boucherit 2011; Rosenhouse 2011), while the majority are associated with a meaning of general futurity.

*yādi

Grammaticalized forms of the active participle *yādi ‘going’ are common future markers throughout much of Morocco and adjacent regions. Reduced and invariant forms often co-exist alongside less grammaticalized reflexes, thereby attesting discrete links in an increasingly advanced grammaticalization chain, as seen in Casablanca *yādī* ~ *ya-*; Caubet 2011.

*māšī

Grammaticalized future markers deriving from the active participle *māšī ‘going’ occur in two distinct geographic pockets, one centered in north-central Morocco and the other in Tunisia and eastern Algeria. In addition to more predictable effects of phonetic erosion and decategorialization, several forms from the latter area display a further example of irregular sound change with the sporadic denasalization of */m/ > /b/, as in Sousse *māš* ~ *bāš* (Talmoudi 1980).

*sāyir

Alone in the sample, Maltese attests a future marker deriving from a grammaticalized form of the active participle *sāyir ‘going’. This can be found in both an inflecting form *sejjer* (SG.M)/*sejra* (SG.F)/*sejrin* (PL) and as the more grammaticalized, invariant forms *ser* and *se* (Vanhove 1993).

2.2.2 Futures from ‘want’ (FUT < WANT)

Grammaticalizations from source constructions indicating desire or volition are another cross-linguistically common origin for future tense markers (Bybee et al.

16 Contact-induced grammaticalization between Arabic dialects

1994; Heine & Kuteva 2002), and are similarly widespread to their FUT < GO counterparts in the Arabic dialect data. Specific grammaticalizations are discussed below.

*yabyā ~ yabyī

Grammaticalized forms of the imperfective verb *yabyā ~ yabyī ‘want’ serve as future markers across a large portion of the Arabic-speaking world, stretching from the Arabian Peninsula across the Red Sea to the greater Sudanic area and then northward through modern Libya. While many Arabic varieties attest only a highly grammaticalized, reduced form of the item (e.g. Abu Dhabi *b-*; Qafisheh 1977), other dialects display direct evidence of multiple stages of the grammaticalization chain, e.g. Ḥarb *yabyā ~ yabā ~ ba-* (Il-Hazmy 1975).

*biddu ~ widdu

Future tense markers arising from grammaticalizations of *biddu ~ widdu ‘want’ are found throughout the broader Levantine area. In their most phonetically reduced forms (e.g. Soukhne *b-*; Behnstedt 1994), they are often superficially indistinguishable from the highly grammaticalized products of *yabyā ~ yabyī discussed above; several dialects, however, provide clear evidence for a distinct chain of development, such as Jebel Ansariye *baddo ~ bado ~ b-* (Lewin 1969) and Cilicia *baddu ~ baddi- ~ bad-* (Procházka 2011). In some varieties, grammaticalizations of *biddu ~ widdu operate alongside other markers of future tense to designate a more specified value: Damascus *baddo ~ b-* is often reported to denote a modal value of possible or planned future, as opposed to the *rāyih-derived forms *rah ~ ha-* which indicate a higher degree of certainty or expectation (cf. Lentin 2011). In other dialects, these forms would appear to have further desemanticized and extended to a value of more general futurity. Future investigation is needed into the degree to which reduced reflexes of *biddu may have merged in mental representation with the continuous aspect marker *b-* present in many of the same varieties; relevant parallels might be drawn with scenarios of near homophony like that found in the dialect of Dhofar, in which continuous *bi-* exists alongside future *bā-* (<*yabyā ~ yabyī; Davey 2016).

*yišā

Thomas Leddy-Cecere

In a number of Yemeni dialects, the future tense marker may be traced to a grammaticalized form of **yīšā* ‘want’. It is notable that in cases such as Sana’a *ša-* this form is used only with the first person singular verb (Watson 1993); in such circumstances, it is possible that its ultimate source should be more properly identified with **ašā* ‘I want’.

**ydawr*

Varieties belonging to the Ḥassāniyya dialect complex of Mauritania and neighboring Mali are recorded as utilizing a grammaticalized form of the verb **ydawr* ‘want’ with a following imperfective verb to denote a value of intentional future. This grammaticalization is relatively “light,” consisting primarily of desemanticization and extension with little in the way of decategorialization or erosion: Nouakchott *ydo:r*, for example, denotes future intent while continuing to operate morphosyntactically as a fully inflected finite verb (Taine-Cheikh 2011).

**byā*

Dialects of southern Morocco and southwestern Algeria occasionally attest grammaticalized forms of **byā* ‘want’ expressing a future tense value. Though lexically similar in origin to the grammaticalizations based on **yabyā* ~ *yibbā* ~ *yibbī* discussed above, the phonological shape of these items (e.g. Marrakech *bya:* ~ *ba-*; Sánchez 2014) recommends an identification of their source in the perfective stem **byā*, which is the typical means for expressing ‘want’ in this area.

2.2.3 Futures from ‘come’ (FUT < COME)

Another cross-linguistically common path of future tense grammaticalization, that involving verbs meaning ‘come, return’ (Bybee et al. 1994; Heine & Kuteva 2002), is represented in the Arabic data by markers originating from a single source etymon, **ʕād* ‘return’.

**ʕād*

Future tense markers traditionally identified as grammaticalized forms of **ʕād* ‘return’ are attested in three locations in the cross-dialectal survey: Yemen, Upper Egypt and interior Tunisia. The forms found in Tunisia and Egypt, Tozeur *ʕa-* and Aswan *ʕa-* (Saada 1984; Schroepfer 2019), are highly

16 *Contact-induced grammaticalization between Arabic dialects*

reduced, and thus difficult to ascribe definitively to a specific source. It is notable that in both of these dialects the markers in question vary with a ‘go’-derived future *ḥa-* and could thus plausibly represent an erosion of the latter in the form of a sporadic lenition of /ḥ/ > /ʕ/ (not to mention that Aswan *ʕa-* on its own might be linked to local *ʕāyiz* ~ *ʕāwiz* ‘want’). At least in the case of the Yemeni forms, however, an origin in *ʕād seems clear, as reduced forms such as Sana’a *ʕā-* display an allomorph *ʕad-* in prevocalic contexts (Watson 1993).

2.2.4 **Futures from purposive constructions (FUT < PURP)**

A further source of future tense markers in the Arabic data involves the grammaticalization of purposive operators. This path is not widely discussed in the cross-linguistic grammaticalization literature, though intriguingly the reverse trajectory, that of PURP < FUT, is noted (Bybee et al. 1994). The primary difficulty would seem to rest in the identification of a clear process of desemanticization, as it is difficult to judge precisely which function between FUT and PURP is more concrete/abstract than the other. Despite this, the occurrence of extension, decategorialization and erosion in the Arabic forms seems to recommend their identification as products of a grammaticalization process.

*ḥattā

Grammaticalizations of *ḥattā ‘in order to’ are used to indicate future tense in areas of northern Mesopotamia, the coastal Levant, and Oman. In terms of geographic distribution and the specific path of phonetic erosion followed, it may be possible to recognize Levantine and Mesopotamian forms like Cypriot Maronite *tta-* and Mosul *də-* (Borg 1985; Jastrow 1979) as representing a single historical innovation, though Oman *ḥa-* ~ *ha-* is more likely an independent development. In the Omani case, the attested use of *ḥa-* with purposive meaning recommends a source in *ḥattā rather than GO-future *rāyih: *šrab ḥa-turwe!* ‘Drink so your thirst be quenched!’ (Reinhardt 1894: 276).

2.2.5 **Futures from ‘to busy oneself with’ (FUT < VERB OF ACTIVITY/PREPARATION)**

A small number of Arabic dialects utilize a future tense marker seeming to derive from a grammaticalized form of a verb meaning ‘to busy oneself’. Such a path

Thomas Leddy-Cecere

of development is not discussed in the cross-linguistic literature on grammaticalization, but perhaps has a counterpart in the use of grammaticalized Southern American English *fixing to* ~ *fixin' a* ~ *fi'na* to express proximate futurity (cf. Wolfram & Schilling-Estes 1998). In any case, obvious desemanticization, extension, decategorialization and erosion of the source form indicate a clear example of grammaticalization in this case.

*lāhī

In the Ḥassāniyya dialects of Mauritania, Mali and far southern Morocco, the future tense marker derives from a grammaticalized form of *lāhī, itself the active participle form of the verb *lha* 'to busy oneself'. Decategorialization is attested in all cases by the lack of adjectival agreement marking predicted for the original participial, and in at least some varieties phonetic erosion is evidenced as well: Mali *lāhi* ~ *lā* (Heath 2003).

2.3 Evidence of replication and diffusion via contact

Of the five grammaticalization paths for Arabic future markers presented above, two merit closer examination in the search for evidence of CIG: those of FUT < GO and FUT < WANT. These paths are identified due to the fact that each is represented in the data by multiple, parallel realizations arising from etymologically distinct but semantically and functionally analogous sources. Such a state of affairs plausibly reflects the result of continued processes of replication, whereby a grammaticalization process occurring in one Arabic variety is transferred to another and recreated using native etymological material.

Both paths identified, however – together representing the great majority of future tense markers attested in the sample – are also extremely common cross-linguistically, and could conceivably have fed multiple independent developments instantiated across the modern Arabic dialect continuum. Key to selecting between an analysis of CIG and one of repeated, internally-motivated grammaticalization is the factor of geography, as in the absence of fine-grained historical sociolinguistic data (see §1.3) this is perhaps the most reliable proxy in positing the feasibility of historical contact between dialects. In the case of CIG, analogous grammaticalization processes ought to be positioned in a geographically contiguous (or near contiguous) bloc, consistent with a history of diffusion via contact between speakers of neighboring dialects. In a scenario of independent development, on the other hand, one should expect the various grammaticalizations to be more or less randomly distributed across the map, equally likely to

16 *Contact-induced grammaticalization between Arabic dialects*

occur in any individual dialect considered.

The geographic incidences of the members of the FUT < GO and FUT < WANT paths both clearly align with the contiguous profile anticipated for the results of CIG. All realizations of FUT < GO future tense markers described connect geographically with other members of the bloc. The large eastern and central zone of *rāyih futures, encompassing southern Mesopotamia, much of the Levant, and the Nile Valley, stretches westward across Libya (where *rāyih-derived forms are recorded alongside *yibbī-based WANT-futures) to include most of Algeria. Directly adjacent to this North African arm of the *rāyih forms are found grammaticalizations of *yādī in Morocco and of *māšī in Tunisia. Further neighboring or co-territorial with the latter two areas are a second set of *māšī-based forms in northern Morocco and the Maltese *sāyir-derived future tense marker, thus completing the connected geographic trend. Future markers representing the path FUT < WANT display a similar spatial contiguity. Grammaticalizations of *biddu ~ widdu in the Levant stretch to reach those of *yabyā ~ yabyī present in the Arabian Peninsula. These in turn span the Red Sea across to the greater Sudanic area and northward through the central Sahara into Libya. Moving to the west and southwest of this zone, the next future markers encountered include grammaticalizations of *byā and *ydawr, respectively. Rounding out the set, forms derived from *yišā exist in close proximity to analogous *yabyā ~ yabyī futures in Yemeni territory. While the integrity of this WANT-future bloc may seem to be challenged by natural features such as the Red Sea and the Sahara Desert, historical and anthropological investigations of the regions in question have rather shown persistent social and cultural connectivity across these would-be barriers (Lydon 2009; Power 2012). This evaluation is supported by the distribution of additional Arabic dialectological isoglosses extending beyond the discussion of CIG.

The geographic contiguity displayed by the representatives of both the FUT < GO and FUT < WANT pathways favors an interpretation of areal diffusion over one of independent, internally motivated occurrence (of the type perhaps evidenced by the more scattered distributions of the sole representatives of FUT < COME and FUT < PURP). The optimal account for the development of the modern Arabic GO- and WANT-futures, together representing the greater part of future tense markers attested in the data, is thus one by which grammaticalization processes leading to the development of new future tense markers have repeatedly been subject to transfer and replication between speakers of neighboring dialects. A CIG-driven analysis such as this has the benefit of accounting for both the development of individual dialect forms and more global trends in source semantics and geo-

Thomas Leddy-Cecere

graphic incidence, and offers a theoretically unified interpretation of the Arabic data obtaining on multiple scales.

3 Conclusion

The analysis summarized above has demonstrated the significant explanatory power of CIG as an account for the development of Arabic future tense markers. Additional proposed occurrences of CIG between Arabic dialects, pertaining to genitive exponents and temporal adverbs meaning ‘now’, are identified and examined in [Leddy-Cecere \(2018\)](#). Together with the future tense data discussed here, these call for corroboration and refinement at the hands of future investigators.

Should further examination provide evidence for a widespread history of CIG between Arabic dialects, this finding could prove instrumental in satisfactorily accounting for a number of so-called “pluriform” developments which have repeatedly vexed students of Arabic dialectology. Defined by Versteegh as functionally analogous but etymologically disparate developments for which “a general trend ... has occurred in all Arabic dialects, and an individual translation of this trend in each area,” dialect contact has most often been dismissed as a causal mechanism for these innovations due to a belief that “typically dialect contact leads to the borrowing of another dialect’s markers, not to the borrowing of a structure that is then filled independently” (1997: 108). CIG provides a theoretical mechanism by which precisely such borrowing and filling may occur, and as such offers the dialectologist a novel analytical tool in the elucidation of structural transfer and diffusion between Arabic varieties.

A critical open question in the application of CIG to the Arabic context, as well as in study of the phenomenon more generally, lies in the problem of agentivity and actuation (as discussed in §1.3). Here, too, further accrual of Arabic data has the potential to inform broader domains of inquiry. If Arabic is established as a productive ground for the study of CIG and significant cases of the historical transfer of grammaticalization pathways between dialects are brought to light, it stands to reason that the same societal and linguistic forces which have motivated these to take place may still be in force, and that observation of synchronic Arabic dialect interaction represents a singular opportunity to catch newly occurring instances of CIG “in the act” and to observe their progress in real time (for at least one such attempt already presented, see [Abuamsha 2016](#)). Studies of this type will enable linguists to add critically lacking synchronic data to their sociolinguistic and psycholinguistic analyses of CIG, and so elaborate

16 *Contact-induced grammaticalization between Arabic dialects*

and strengthen ongoing theorizations of a revelatory new dimension of contact-induced language change.

Further reading

For a complete theoretical discussion of contact-induced grammaticalization, see Heine & Kuteva (2003; 2005). The former work is an article-length sketch of the proposal and is valuable as a direct and concise reference, while the latter provides a more elaborated description with additional linguistic examples. Matras (2009) provides valuable commentary and critique of Heine & Kuteva's work while simultaneously extending exploration to the psycholinguistic and sociolinguistic dimensions of CIG.

For an overview of grammaticalization processes in the development of Arabic future tense markers (though without reference to contact) see Stewart (1998). A more detailed treatment of CIG processes in the development of the Arabic future markers and other morphosyntactic features may be found in Ledy-Cecere (2018).

Acknowledgements

I am grateful to Drs. Kristen Brustad and Daniel Law for their contribution to the broader work which this chapter reflects, as well to the participants and organizers of the workshop "Arabic and Contact-Induced Change" at the 23rd International Conference on Historical Linguistics for their insightful feedback relating to the ideas discussed here.

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1110007. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Thomas Leddy-Cecere

Abbreviations

CIG	contact induced	M	masculine
	grammaticalization	PL	plural
F	feminine	PURP	purposive
FUT	future	RL	recipient language
invar.	invariant	SG	singular
		SL	source language

References

- Abdel-Massih, Ernest, Zaki Abdel-Malek, El-Said Badawi & Ernest McCarus. 2009. *A reference grammar of Egyptian Arabic*. Washington, D.C.: Georgetown University Press.
- Abuamsha, Duaa. 2016. The future marker in palestinian arabic: an internal or contact-induced change? In Lindsay Hracs (ed.), *Proceedings of the 2016 annual conference of the canadian linguistic association*.
- Andersen, Henning. 2008. Grammaticalization in a speaker-oriented theory of change. In Thórhallur Eythórsson (ed.), *Grammatical change and linguistic theory: The Rosenthal Papers*, 11–44. Amsterdam, Philadelphia: John Benjamins.
- Behnstedt, Peter. 1994. *Der arabische Dialekt von Soukhne (Syrien). Teil 2: Phonologie, Morphologie, Syntax. Teil 3: Glossar*. Wiesbaden: Harrassowitz.
- Behnstedt, Peter & Manfred Woidich. 2005. *Arabische Dialektgeographie: Eine Einführung*. Leiden: Brill.
- Bisang, Walter. 1998. Grammaticalization and language contact, constructions and positions. In Ana Giacalone-Ramat & Paul Hopper (eds.), *The limits of grammaticalization*, 13–58. Amsterdam, Philadelphia: John Benjamins.
- Borg, Alexander. 1985. *Cypriot Arabic: a historical and comparative investigation into the phonology and morphology of the arabic vernacular spoken by the Maronites of Kormakiti village in the Kyrenia district of north-west Cyprus*. Stuttgart: Deutsche Morgenländische Gesellschaft.
- Boucherit, Aziza. 2011. Algiers Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Bybee, Joan. 2003. Cognitive processes in grammaticalization. In Michael Tomasello (ed.), *The new psychology of language*, vol. 2, 145–167. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago, London: Chicago University Press.

16 Contact-induced grammaticalization between Arabic dialects

- Caubet, Dominique. 2011. Moroccan Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Claudi, Ulrike & Bernd Heine. 1986. On the metaphorical basis of grammar. *Studies in Language* 10(2). 297–335.
- Dahl, Östen. 2001. Principles of areal typology. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, 1456–1470. Berlin: Walter de Gruyter.
- Davey, Richard J. 2016. *Coastal Dhofari Arabic: A sketch grammar*. Leiden: Brill.
- Davies, Humphrey. 1981. *17th-century Egyptian arabic: a profile of the colloquial material in yūsuf al-Širbīnī's hazz al-quhūf fī šarḥ qaṣīd abī Šādūf*. Berkeley: University of California doctoral dissertation.
- Elias, Elias & Edward Elias. 1981. *Egyptian Arabic manual for self study*. 19th ed. Cairo: Elias Modern Publishing House.
- Epps, Patience. 2005. Areal diffusion and the development of evidentiality: Evidence from hup. *Studies in Language* 29(3). 617–650.
- Grosjean, François. 2001. The bilingual's language modes. In Janet Nicol (ed.), *One mind, two languages: Bilingual language processing*, 1–22. Malden: Blackwell.
- Haase, Martin. 1992. *Sprachkontakt und Sprachwandel im Baskenland: die Einflüsse des Gaskognischen und Französischen auf das Baskische*. Hamburg: Buske.
- Hansen, Maj-Britt Mosegaard & Jacqueline Visconti. 2009. On the diachrony of “reinforced” negation in French and Italian. In Corinne Rossari, Claudia Ricci & Adriana Spiridon (eds.), *Grammaticalization and pragmatics: Facts, approaches, theoretical issues*, 137–171. Bingley: Emerald Group.
- Haspelmath, Martin. 1998. Does grammaticalization need reanalysis? *Studies in Language* 22(2). 271–287.
- Heath, Jeffrey. 2003. *Hassaniya Arabic (Mali): Poetic and ethnographic texts*. Wiesbaden: Harrassowitz.
- Heine, Bernd. 2004. Grammaticalization. In Brian Joseph & Richard Janda (eds.), *The handbook of historical linguistics*, 575–601. Oxford: Wiley-Blackwell.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3). 529–572.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.

Thomas Leddy-Cecere

- Hopper, Paul & Elizabeth Traugott. 2003. *Grammaticalization*. Second edition. Cambridge: Cambridge University Press.
- Il-Hazmy, Alayan Mohammed. 1975. *A critical and comparative study of the spoken dialect of the Ḥarb tribe in Saudi Arabia*. Leeds: University of Leeds doctoral dissertation.
- Jastrow, Otto. 1979. Zur arabischen Mundart von mossul. *Zeitschrift für Arabische Linguistik* 2. 76–99.
- Law, Daniel. 2014. *Language contact, inherited similarity and social difference: The story of linguistic interaction in the Maya Lowlands*. Amsterdam, Philadelphia: John Benjamins.
- Le Page, R. B. & Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- Leddy-Cecere, Thomas. 2018. *Contact-induced grammaticalization as an impetus for Arabic dialect development*. Austin: University of Texas doctoral dissertation.
- Lehmann, Christian. 1985. Grammaticalization: synchronic variation and diachronic change. *Lingua e Stile* 20. 303–318.
- Lehmann, Christian. 2015. *Thoughts on grammaticalization*. Berlin: Language Science Press.
- Lentin, Jérôme. 2011. Damascus Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Lewin, Bernhard. 1969. *Notes on Cabali: The Arabic dialect spoken by the Alawis of “Jebel Ansariye.”* Stockholm: Almqvist & Wiksell.
- Lydon, Ghislaine. 2009. *On trans-Saharan trails: Islamic law, trade networks, and cross-cultural exchange in nineteenth-century western Africa*. Cambridge: Cambridge University Press.
- Mahdi, Qasim R. 1985. *The spoken Arabic of baṣra, iraq: a descriptive study of phonology, morphology and syntax*. Exeter: University of Exeter doctoral dissertation.
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2011. Grammaticalisation and language contact. In Bernd Heine & Heiko Narrog (eds.), *Handbook of grammaticalisation*, 279–290. Oxford: Oxford University Press.
- Miller, Catherine, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.). 2007. *Arabic in the city: Issues in dialect contact and language variation*. London: Routledge.
- Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.

16 *Contact-induced grammaticalization between Arabic dialects*

- Power, Timothy. 2012. *The Red Sea from Byzantium to the Caliphate, AD 500–1000*. Cairo: The American University in Cairo Press.
- Procházka, Stephan. 2011. Cilician Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Qafisheh, Hamdi. 1977. *A short reference grammar of Gulf Arabic*. Tucson, Arizona: University of Arizona Press.
- Ramat, Paolo. 2011. Adverbial grammaticalization. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of grammaticalization*, 502–510. Oxford: Oxford University Press.
- Reinhardt, Carl. 1894. *Ein arabischer Dialekt gesprochen in ‘Omān und Zanzibar*. Stuttgart, Berlin: W. Spemann.
- Rosenhouse, Judith. 2011. Jerusalem Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Ross, Malcolm. 2007. Calquing and metatypy. *Journal of Language Contact* 1(1). 116–143.
- Rubin, Aaron D. 2005. *Studies in Semitic grammaticalization*. Winona Lake, IN: Eisenbrauns.
- Saada, Lucienne. 1984. *Eléments de description du parler Arabe de Tozeur (Tunisie): Phonologie, morphologie, syntaxe*. Paris: Librairie Orientaliste Paul Geuthner.
- Sánchez, Pablo. 2014. *El árabe vernáculo de Marrakech: Análisis lingüístico de un corpus representativo*. Zaragoza: Prensas de la Universidad de Zaragoza.
- Schroepfer, Jason. 2019. Notes on aswan arabic. Unpublished manuscript, University of Texas.
- Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23. 193–229.
- Stewart, Devin. 1998. Clitic reduction in the formation of modal prefixes in the post-classical arabic dialects and classical arabic “sa-/sawfa”. *Arabica* 45(1). 104–128.
- Taine-Cheikh, Catherine. 2011. Ḥassānīyya Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Talmoudi, Fathi. 1980. *The Arabic dialect of Sūsa (Tunisia)*. Göteborg: Acta Universitatis Gothoburgensis.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Vanhove, Martine. 1993. *La langue maltaise: etudes syntaxiques d’un dialecte arabe «périphérique»*. Wiesbaden: Harrassowitz.

Thomas Leddy-Cecere

- Versteegh, Kees. 1997. *The Arabic language*. First edition. Edinburgh: Edinburgh University Press.
- Vollers, Karl. 1895. *The modern Egyptian dialect of Arabic: A grammar with exercises, reading lessons and glossaries. translated by francis burkitt*. Cambridge: Cambridge University Press.
- Watson, Janet C. E. 1993. *A syntax of Ṣanʿānī Arabic*. Wiesbaden: Harrassowitz.
- Wolfram, Walt & Natalie Schilling-Estes. 1998. *American English: Dialects and variation*. Malden: Blackwell.

Chapter 17

Contact and calquing

Stefano Manfredi

CNRS, SeDyL

The notion of calquing refers to the transfer of semantic and syntactic patterns deprived of morphophonological matter. By providing examples of lexical and grammatical calques in a number of Arabic dialects and Arabic-based contact languages, this chapter identifies ways to relate the process of calquing to Van Coetsem’s psycholinguistic principle of language dominance.

1 Introduction

In its simplest definition, CALQUING is a type of contact-induced change in which a word or sentence structure is transferred without actual morphemes (Thomason 2001: 260). Calques are sometimes called loan translations as they typically represent a word-by-word (or morpheme-by-morpheme) translation of a lexeme or a sentence from another language. Heath (1984: 367) labels this process “pattern transfer” and distinguishes it from “matter borrowing” which is instead linked to the integration of morphophonological material. Ross (2007), for his part, points out that that calquing can also have important grammatical effects, and he considers it a necessary precondition for contact-induced morphosyntactic restructuring (what Ross calls “metatypy”).

Broadly speaking, we can distinguish two types of calquing: lexical calquing, which entails the transfer of semantic properties of lexical items, and grammatical calquing, which instead implies the transfer of the functional properties of morphemes and syntactic constructions. Using Ross’s words (2007: 126), lexical calquing consists of remodelling lexical “ways of saying things”, whereas grammatical calquing consists of remodelling grammatical “ways of saying things”. Despite this fundamental difference, lexical and grammatical calquing share a



Stefano Manfredi

single cause: bilingual speakers' need to express the same meaning in two languages (Sasse 1992: 32). This also means that everything that expresses meaning (i.e. morphemes, lexemes, and constructions) can, in principle, be a source of calquing.

Focusing mainly on the transfer of linguistic matter, Van Coetsem (1988) does not overtly mention the possibility of transferring lexical and grammatical meanings through calquing. This chapter thus aims at relating contact induced changes produced by calquing to the principle of language dominance as postulated by Van Coetsem.

2 Contact-induced changes and calquing

2.1 Lexical calquing

According to Haspelmath (2009: 39), a lexical calque is a lexical unit that was created by an item-by-item translation of the source unit. This type of contact-induced change occurs as bilingual speakers reorganise the lexicon of one of their languages to match the semantic organisation of the other (Ross 2007: 132). Adopting the psycholinguistic standpoint of language dominance, Winford (2003: 345) regards lexical calquing as a subtype of lexical borrowing, which is a combination of recipient language (RL) lexemes in imitation of source language (SL) semantic patterns. In contrast, I will show that, though lexical calquing can easily be triggered by RL-dominant speakers, it can also be a product of imposition via SL agentivity. In order to do this, I will mainly focus on calquing of compound nouns. A compound noun is here defined as a series of two or more lexemes, which is semantically conceived as a single unit. Each component of the compound can function as a lexeme independent from the other(s), and may show some phonological and/or morphological constraints within the compound when compared to its isolated syntactic usage (Bauer 2001). Against this backdrop, I will specifically discuss noun–noun compounds as they represent the more uniform phenomenon of nominal compounding in the world's languages (Pepper forthcoming). As we will see, the transfer of the semantics of compound nouns does not imply any morphosyntactic change in Arabic, as calqued compounds are typically adjusted to fit RL morphosyntactic patterns.

Generally speaking, lexical calquing through borrowing can occur in indirect contact situations characterized by a very low degree of bilingualism. This is because RL monolinguals can also be agents of lexical borrowing (Van Coetsem 1988: 10) Typical instances of lexical calquing via RL agentivity are related to the

transfer of the semantic patterns of English compound nouns in modern Arabic dialects. This kind of transfer is linked to the expansion of the non-core Arabic lexicon for expressing previously unknown concepts. A prime example is the English calque *lōhit il-maḡatīh* ‘keyboard’ (lit. ‘the board of keys’) in Egyptian Arabic (Wilmsen & Woidich 2011: 9). Here, it can be clearly seen that the transfer of the semantic organization of the SL compound noun does not affect the morphosyntax of the RL, as the word order of the English nominal juxtaposition is reversed to fit the Arabic construct state.

Lexical calquing can also take place in prolonged contact situations, as testified by numerous Italian compounds in Maltese. A singular case of mixed calquing is that of *wicc tost* ‘shameless person’ (lit. ‘tough face’) deriving from the Italian compound *faccia tosta* ‘shameless person’ (lit. ‘tough face’) (Aquilina 1987). On the one hand, the first lexical item of the compound presents an Arabic phonological form while expressing semantic properties associated with the lexeme ‘face’ in Italian. On the other hand, the second lexical item clearly results from the borrowing of the adjective *tosto* ‘hard, tough’ retaining both the Italian phonological matter and semantic properties. The mixed nature of this compound brings to the fore the complementary relationship between RL and SL agentivity and shows that it is not always a trivial matter to distinguish between imposition and borrowing. However, Maltese also gives evidence of genitive compounds in which both lexical components have an Arabic phonological form coupled with Italian semantic properties. This is the case of the compound nouns *saba’ ta’ sieq* ‘toe’ calqued on the Italian *dito del piede* ‘toe’ (Pepper forthcoming). Such instances of lexical calquing clearly mirror semantic properties of SL lexemes and they most plausibly result from borrowing via RL agentivity (cf. Lucas & Čeplö, this volume).

Ḥassāniyya Arabic, for its part, presents many compound nouns that are traditionally analysed in terms of substratum interference from Zenaga Berber (Taine-Cheikh 2008; 2012). Also in this case, the transfer of the semantic properties of the SL does not produce any morphosyntactic change in Arabic, as we can see in the pairs of examples in (1) and (2).

- (1) a. Ḥassāniyya Arabic (Taine-Cheikh 2008: 126)
 kṛaʔ lə-ɣṛab
 foot DEF-crow
 ‘aquatic herbaceous plant’ (lit. ‘crow’s foot’)
- b. Zenaga Berber (Taine-Cheikh 2008: 126)

Stefano Manfredi

aḏaʔrən tayyal
 foot GEN crow
 ‘aquatic herbaceous plant’ (lit. ‘crow’s foot’)

- (2) a. Ḥassāniyya Arabic (Taine-Cheikh 2008: 126)
 sayllāl lə-ʕrāgib
 ripper DEF-ankle.PL
 ‘honey badger’ (lit. ‘ripper of ankles’)
- b. Zenaga Berber (Taine-Cheikh 2008: 126)
 aməssāf ən ūrʒan
 ripper GEN ankle.PL
 ‘honey badger’ (lit. ‘ripper of ankles’)

Taine-Cheikh (2008: 126) stresses that it is somewhat difficult to trace back the origin of these compounds. Accordingly, she speaks of a process of convergence between the two languages, rather than determining the direction of the semantic transfer. However, it should be observed that these compound nouns are not attested in other spoken varieties of Arabic. Furthermore, since at least the mid-twentieth century, Berbers in Mauritania have been gradually losing competence in Zenaga, in favour of Arabic (Taine-Cheikh 2012: 100), while Zenaga is rarely acquired as second language by Ḥassāniyya Arabic speakers. In such a context, the most probable agents of contact-induced change were former Berber-dominant speakers who gradually shifted to Arabic. Thus, it seems plausible that the transfer of the semantic properties of Zenaga compounds has been achieved through imposition, rather than through borrowing.

Nigerian Arabic also shows interesting instances of lexical calquing as a consequence of a longstanding contact with Kanuri, a Nilo-Saharan language widely spoken in the Lake Chad area. Owens (2015; 2016) gives evidence of the transfer of the semantic properties of numerous compound nouns including the lexeme *rās* ‘head’. Similar to the previous instances of compound calquing, the integration of Kanuri semantic patterns does not affect the Arabic morphosyntax, as we can see in the pairs of examples in (3) and (4).

- (3) a. Nigerian Arabic (Owens 2016: 69)
 rās al-bēt
 head DEF-house
 ‘roof’ (lit. ‘head of house’)
- b. Kanuri (Owens 2016: 69)

kəla fəto-be
 head house-GEN
 ‘roof’ (lit. ‘head of house’)

- (4) a. Nigerian Arabic (Owens 2016: 65)

ɾās al-qalla
 head DEF-corn
 ‘tassel’ (lit. ‘head of corn’)

- b. Kanuri (Owens 2016: 65)

kəla argəm-be
 head corn-GEN
 ‘tassel’ (lit. ‘head of corn’)

According to Owens (2016: 65), Kanuri–Arabic bilingualism, with Arabic being a minority language, would have been the foremost factor underlying the transfer of these compound nouns into Nigerian Arabic. He further stresses that Kanuri is the main source of compound nouns in a number of other minority languages in the area (e.g. Kotoko, Glayda, and Fulfulde) and that there is little evidence of Kanuri to Arabic shift in the region (Owens 2014: 147). However, the fact that Kanuri represents the majority language of northeastern Nigeria, does not shed light on the transfer mechanism lying behind lexical calquing in Nigerian Arabic. This is because speakers can be linguistically dominant in a socially subordinate language (Winford 2005: 376). In fact, such contact settings are closely tied to SL agentivity, as the youngest bilingual generations tend to impose semantic features from their dominant language (i.e. Kanuri) onto the ancestral language (i.e. Arabic). It is only at a later stage that these innovations are borrowed by older bilingual speakers who are still dominant in Arabic.

The fact that Nigerian Arabic speakers have gradually developed a high bilingual proficiency in Kanuri is also testified by the transfer of a number of idiomatic expressions. In this regard, Ross (2007: 122) observes that calquing of meaning is not only reflected in word compounding, but also in lexical collocations of idiomatic expressions. These are combinations of lexical items that are semantically idiosyncratic as they have a pairing of form and meaning that cannot be predicted from the rest of the grammar. The pair of examples in (5) provide evidence of an idiomatic Kanuri calque in Nigerian Arabic.

- (5) a. Nigerian Arabic (Ritt-Benmimoun et al. 2017: 77)

Stefano Manfredi

šuqul šāl rās-i
 something carry.PRF.3SG.M head-OBL.1SG
 ‘Something distracted me.’ (Lit. ‘Something carried my head.’)

- b. Kanuri (Ritt-Benmimoun et al. 2017: 77)

awode kəla gō-zə-na
 something head carry-3SG-PRF
 ‘Something distracted me.’ (Lit. ‘Something carried head.’)

Given that idiomatic expressions are syntactically compositional (i.e. their lexical components behave syntactically as they do in non-idiomatic expressions), it is not only the meanings expressed by the lexeme ‘head’ which correspond between Nigerian Arabic and Kanuri, but also their idiomatic collocations, which align between the two languages (Owens 2014: 157). Besides, it is worth noting that also idiomatic expressions are adjusted to fit RL morphosyntactic patterns. This is evidenced by the inalienable possession of body parts in Nigerian Arabic (*rās-i* ‘my head’), which is instead unattested in the SL (*kəla* ‘head’). Even if we cannot exclude the possibility that these kinds of calques are a product of borrowing, it is evident that their integration needs a high proficiency in the SL for individuating the single idiomatic collocations of lexical items. Furthermore, differently from borrowed calques, imposed idiomatic expressions can significantly affect the lexical semantics of the RL created by SL-dominant bilinguals, and thus produce grammatical changes in the long run.

Finally, lexical calquing via SL-agentivity can also take place in extreme contact situations such as creolization. For instance, Juba Arabic, the Arabic-based pidgincreole spoken in South Sudan, shows numerous calques in which Arabic-derived lexemes are compounded according to the semantic patterns of Bari, the main substrate language of Juba Arabic (Nakao 2012; Manfredi 2017: 50; Avram, this volume). As we can see in (6) and (7), the word order in Juba Arabic compounds follows the order of Bari compounds. However, this cannot be seen as an innovative morphosyntactic development, as the possessed–possessor order matches also with the Arabic lexifier.

- (6) a. Juba Arabic (Nakao 2012: 136)

éna ta séjera
 eye GEN tree
 ‘fruit’ (lit. ‘eye of tree’)

- b. Bari (Nakao 2012: 136)

koŋe lo-ködini
 eye GEN-tree
 ‘fruit’ (lit. ‘eye of tree’)

- (7) a. Juba Arabic (Nakao 2012: 137)
 ida ta fil
 hand GEN elephant
 ‘trunk’ (lit. ‘hand of elephant’)
- b. Bari (Nakao 2012: 137)
 könin lo-tome
 hand GEN-elephant
 ‘trunk’ (lit. ‘hand of elephant’)

Given that the asymmetric contact situation leading to creole formation limits access to the superstrate language (i.e. Sudanese Arabic), the semantic patterns of substrate languages (i.e. Bari) can be easily carried over into the creole in ways peculiar to imposition via SL-agentivity.

All things considered, unlike lexical borrowing, lexical calquing allows for a semantic overlapping of RL and SL lexical entries and it can also produce important structural changes.

2.2 Grammatical calquing

Grammatical calquing brings about a match between the grammatical categories of two languages and the memberships of these categories (Ross 2007: 132). Heine & Kuteva (2005) suggest that the grammatical changes induced by calquing can be better analysed in terms of contact-induced grammaticalization (see also Leddy-Cecere, this volume). In fact, the calquing of the semantic properties of lexical and grammatical items may lead to the grammaticalization of innovative syntactic structures in the RL matching with those of the SL. From the traditional socio-historical perspective of contact-induced change (Thomason & Kaufman 1988), grammatical calquing is basically seen as a product of language shift. In contrast, Ross (2007: 131) argues that grammatical calques can widely occur in situations of language maintenance. Actually, the different grammatical outputs of calquing mainly depend on the way in which they are transferred from the SL into the RL and, by extension, on different kinds and degrees of bilingualism.

For the purposes of this chapter, I distinguish between three different types of grammatical calquing:

Stefano Manfredi

- Calquing of polyfunctionality of lexical items without syntactic change;
- Calquing of polyfunctionality of grammatical items leading to syntactic change;
- Narrow syntactic calquing (without calquing of polyfunctionality of lexical/grammatical items).

Being lexical in nature, the first of these three types of grammatical calquing can be triggered by both imposition via SL agentivity and borrowing via RL agentivity, whereas the two latter types are likely to result only from imposition via SL agentivity.

Calquing of polyfunctionality patterns of lexical items is by far the most common type of grammatical calquing, and it can be exemplified by the comparison of reflexive anaphors in different Arabic dialects. As is well known, Classical and Standard Arabic express a reflexive meaning either by means of agent-oriented derived verbs lacking an overtly expressed patient (e.g. *istaḥamma* ‘he washed himself’) or by anaphoric constructions in which the syntagm *naḥs*-PRO.POSS ‘soul-PRO.POSS’ marks coreferentiality between the agent and the patient of the predicate (e.g. *qatala naḥsa-hu* ‘he killed himself’). Nevertheless, as a result of contact with different languages, a number of modern Arabic dialects have grammaticalized other lexical sources for expressing a reflexive meaning. Western Maghrebi dialects are a case in point. As we can see in (8)–(9), both Moroccan and Ḥassāniyya Arabic have grammaticalized the nominal syntagm *raḥḥ*=PRO.POSS ‘head-PRO.POSS’ as default reflexive anaphor.

- (8) Moroccan Arabic (D. Caubet, personal communication)

qtəl rās-o

kill.PRF.3SG.M head-3SG.M

‘He killed himself.’ (Lit. ‘He killed his head.’)

- (9) Ḥassāniyya Arabic (Taine-Cheikh 2008: 16)

ktəl rāḥ-u

kill.PRF.3SG.M head-3SG.M

‘He killed himself.’ (Lit. ‘He killed his head.’)

This reflexive use of the lexeme ‘head’ has generally been interpreted as substrate interference from Berber languages (El Aissati 2011: 197), in which the same grammaticalization path is attested, as shown in the following examples from Tarifit and Zenaga:

- (10) Tarifit Berber (Kossmann 2000: 95)
 yəṭšaθ iħəf nnəs
 beat.PRF.3SG.M head POSS.3SG.M
 ‘He beats himself.’ (Lit. ‘He beats his head.’)
- (11) Zenaga Berber (Taine-Cheikh 2008: 126)
 yəʔna iʔf-ən-š
 kill.PRF.3SG.M head-GEN-3SG.M
 ‘He killed himself.’ (Lit. ‘He killed his head.’)

The lexeme for ‘head’ is the second most common source of grammaticalization of reflexive anaphors worldwide (König et al. 2013) and its occurrence is particularly common in West Africa (Heine 2011: 50). In this scenario, it should be stressed that the reflexive function of the lexeme ‘head’ is an innovative feature of both Arabic and Berber varieties of northwestern Africa. Other Berber languages typically use the reflexive anaphor *iman*-POSS ‘soul-POSS’, as we can see in the following example from Kabyle.

- (12) Kabyle (Mettouchi 2012)
 n-səlk-dd iman-ntəy
 1PL-spare.PRF-PROX soul.ABS.SG.M-POSS.1PL.F
 ‘We saved ourselves.’

In addition, the known Arabic–Berber contact situation, in which second language learners of Berber only played a marginal role in triggering contact-induced change in Arabic, suggests that the contact induced grammaticalization of ‘head’ in westernmost Arabic dialects resulted from an imposition enacted by former Berber-dominant speakers.

A similar instance of calquing in the domain of anaphoric reflexive constructions is found in Kordofanian Baggara Arabic, a western Sudanic dialect spoken in the Nuba Mountains area, in central Sudan. In this case, the source of the reflexive anaphor is the lexeme for ‘neck’, as we can see in (13).

- (13) Kordofanian Baggara Arabic (Manfredi 2010: 176)
 abrahīm gaṣṣa ragabt-a
 Ibrahim cut.PRF.3SG.M neck-3SG.M
 ‘Ibrahim cut himself.’ (Lit. ‘Ibrahim cut his neck.’)

Different from ‘head’, the grammaticalization of ‘neck’ as a reflexive anaphor is quite rare in Africa (Heine 2011: 50), but it is attested in a number of Niger-

Stefano Manfredi

Kordofanian languages spoken in the same region. Such is the case of Tagoi (14) and Koalib (15).

- (14) Tagoi (Alamin 2015: 26)
 t-ágám t-ùrún inní
 NC-neck NC-POSS.3 kill.PRF.3
 ‘He killed himself.’ (Lit. ‘He killed his neck.’)

- (15) Koalib (N. Quint, personal communication)
 èɽnyé r-ókwɛ̀ r-ùɣwún
 kill NC-neck NC-POSS.3
 ‘to kill oneself’ (lit. ‘to kill one’s neck’)

Similarly to the situation described with reference to western Maghrebi dialects, Arabic-speaking groups in the Nuba Mountains have hardly developed any bilingual competence in local Niger-Kordofanian languages. Therefore, it seems likely that the calquing of the polyfunctionality patterns of ‘neck’ has been imposed by Arabized populations who were dominant in the SL.

Maltese also provides remarkable examples of calquing of polyfunctionality of lexical items. This is particularly evident in the domain of auxiliary verbs (Van-hove 1993; Vanhove et al. 2009). A well-known example is that of the lexical verb *gie* ‘come’ used as an auxiliary for expressing a dynamic passive (16) in the same way as Italian (17).

- (16) Maltese (Borg & Azzopardi-Alexander 1997: 214)
 it-tabib gie afdat bi-l-każ
 DEF-doctor come.PRF.3SG.M trust.PTCP.PASS with-DEF-case
 ‘The doctor was entrusted with the case.’ (Lit. ‘The doctor came entrusted with the case.’)

- (17) Italian (own knowledge)
 non venne creduto
 NEG come.PAST.3SG.M trust.PTCP.PAST
 ‘He was not trusted.’ (Lit. ‘He did not come trusted.’)

Even if imposition played a role in the emergence of Maltese (Lucas & Čeplo, this volume), it is generally accepted that intertwined languages emerge mainly from a widespread process of borrowing in Van Coetsem’s terminology (Winford 2005: 397; Manfredi 2018). This suggests that, unlike the aforementioned

grammaticalization of reflexive anaphors in Arabic dialects, the calquing of poly-functionality of lexical verb ‘come’ in Maltese was most likely triggered by agentivity of RL dominant speakers.

Regardless of the different contact situations, what holds all the previous instances of grammatical calquing together is the fact that the transfer of patterns of grammaticalization did not produce any syntactic change in Arabic. In contrast to the above, the calquing of polyfunctionality of grammatical items can be accompanied by important typological changes. This is the case of the grammaticalization of prototypical passive constructions in Juba Arabic (Manfredi 2017: 92; 2018: 415). As we can see in (18), the South Sudanese pidgincreole presents an innovative passive construction in which the patient occupies the syntactic slot of a preverbal subject, whereas the oblique-marked agent is introduced by the comitative preposition *ma-* ‘with’.

- (18) Juba Arabic (Manfredi 2017: 86)
 bab de kasurú ma-jón
 door PROX.SG break.PASS with-John
 ‘This door has been broken by John.’ (Lit. ‘This door has been broken with John.’)

Interestingly, this prototypical passive construction is not attested in the lexifier language of Juba Arabic (i.e. Sudanese Arabic), which instead makes use of impersonal passive constructions with a default 3PL.M subject.

- (19) Sudanese Arabic (own knowledge)
 kassaru-hu
 break.PRF.3PL.M-3SG.M
 ‘It got broken.’ (Lit. ‘They have broken it.’)

Indeed, the grammaticalization of this complex syntactic structure is the result of the calquing of the functional properties associated with the comitative preposition of the main substrate language, Bari. Bari presents the same kind of prototypical passive construction in which an oblique-marked agent is introduced by the preposition *ko-* ‘with’.

- (20) Bari (Owen 1909: 65)
 niena wuret a-wur-ö ko-nan
 PROX.SG book 3SG.PAST-write-PASS with-1SG
 ‘This book has been written by me.’ (Lit. ‘This book has been written with me.’)

Stefano Manfredi

If we assume that the emergence of creole languages is always induced by the disruption of the transmission of the lexifier language (Comrie 2011), we can conclude that Bari speakers have imposed the semantics of their dominant language on a grammatical item derived from Arabic, and thus induced profound changes in the word order of the creole when compared to its lexifier language.¹ In light of the above, the contact dynamics lying behind the calquing of polyfunctionality of grammatical items are quite restrictive as they are most likely a product of imposition via SL agentivity.

The third kind of grammatical calquing is linked to the transfer of syntactic patterns without transfer of polyfunctionality of either lexical or grammatical items. This narrow type of syntactic calquing can be exemplified by possessor doubling in Central Asian Arabic (Ratcliffe 2005). Clitic doubling is a construction in which a clitic co-occurs with a full nominal phrase in argument position, forming a discontinuous constituent with it. Various forms of clitic doubling have arisen in a number of Arabic varieties as a result of contact with different substrate/adstrate languages (Souag 2017). In regard to possessive constructions, Arabic typically presents a possessed–possessor order. In contrast, Central Asian Arabic (21) gives evidence of the opposite order with obligatory possessor doubling in the same way as Tajik (22).

- (21) Central Asian Arabic (Ratcliffe 2005; Souag 2017: 56)

amīr wald-u
prince son-3SG.M
'the prince's son'

- (22) Tajik (Souag 2017: 56)

buħoro universitet-aš
Bukhara university-3SG
'Bukhara University'

Souag (2017: 157) states that double possessor constructions in Central Asian Arabic are instances of grammatical calquing, accommodated through the reinterpretation of pre-existing topicalized constructions. This means that, unlike the syntactic changes induced by the calquing of polyfunctionality of morphemes, the emergence of double possessor constructions in Bukhara Arabic would have been favoured by a formal congruence between SL and RL syntactic structures. As such, this instance of contact-induced morphosyntactic restructuring (i.e. metatypy)

¹This kind of syntactic change accompanied by the calquing of semantic properties of substrate items in creole languages is traditionally labelled "relexification" (Lefebvre 1998).

does not derive from a direct copying of a double possessor construction. Rather, it consists in speakers expressing a possessive meaning in Arabic by using a construction which they equate with the construction in adstratal languages (Ross 2007: 128). If we consider that the youngest speakers of Central Asian Arabic are gradually losing competence in their ancestral language in favour of socially dominant languages (Chikovani 2005: 128), it is plausible to think that such kind of syntactic restructuring can only be a result of imposition via SL-agentivity. Still, given our limited diachronic knowledge, we cannot exclude the hypothesis of an early process of borrowing enacted by former Arabic-dominant speakers.

3 Conclusions

Van Coetsem (1988: 20) suggests that the variable outcomes of language contact are primarily a reflex of the “stability gradient” of language, which induces speakers to preserve the domains of their dominant language that are less affected by change. As lexicon is the most unstable linguistic domain, it is likely to be transferred via RL-agentivity. In contrast, morphosyntax and phonology are considered to be relatively stable domains and they are expected to be transferred only via SL-agentivity. Against this background, it is unclear how the transfer of semantic features deprived of morphophonological matter should be understood in relation to the linguistic dominance of the agents of contact-induced change.

If we look at the previously analysed instances of lexical calquing (§2.1), it is evident that the transfer of the semantic features of nominal compounds can take place within speech communities with a very low degree of bilingualism, as in the case of Egyptian-Arabic-dominant speakers borrowing the semantics of English compounds. But it is also true that compound calquing can be a product of imposition resulting from ongoing language shift or pidginization, and the transfer of semantic features of single lexical items within idiomatic expressions always requires a widespread proficiency in the SL, as in the case of Arabic–Kanuri bilingualism in northern Nigeria.

As far as grammatical calquing is concerned (§2.2), I have shown that calquing of the polyfunctionality of lexical items can be triggered either by imposition, as in the case of substrate interference in Ḥassāniyya and Baggara dialects, or by borrowing in the emergence of intertwined languages such as Maltese. Calquing of polyfunctionality of grammatical items, for its part, requires a higher degree of linguistic abstraction for the identification of a functional overlap between morphemes. Accordingly, this type of transfer will typically occur via imposition by SL-dominant speakers in deep contact situations such as creolization. In the

Stefano Manfredi

same manner, narrow syntactic calquing requires high bilingual proficiency, as it necessitates the recognition of some formal congruence between the SL and the RL, as shown by the emergence of possessor doubling in Central Asian Arabic.

To stay somewhat in line with the stability gradient principle, we could argue that, in absence of the transfer of linguistic matter, the semantic properties of morphemes and syntactic constructions are more stable than those of lexical items. However, such a generalization would be misleading without an in-depth knowledge of the sociolinguistic circumstances underlying a specific instance of second language acquisition (i.e. symmetric bilingualism, asymmetric bilingualism, multilingualism, pidginization/creolization). Thus, it becomes evident that the recognition of different patterns of bilingualism within the same community remains the only way to identify the transfer type at play in a given contact situation, regardless of its different structural outputs.

Drawing on the available literature, this chapter has surveyed only a few instances of calquing induced by contact between Arabic and other languages. This is mainly because we lack information about calquing in dialect contact situations. Indeed, it is regrettable that studies dealing with dialect contact and new dialect formation are still exclusively focused on the diffusion of few lexical and morphophonological features, while disregarding the transfer of semantic and syntactic patterns. Fine-grained analyses of semantic changes induced by dialect contact thus remain a major desideratum for the development an aggregate variationist Arabic dialectology.

Further reading

Keesing (1988) adopts the notion of calquing and describes the transfer of semantic properties of Oceanic morphemes in Melanesian Pidgin.

Meyerhoff (2009), by focusing on the notions of replication, transfer, and calquing, strengthens connections between variationist sociolinguistics and contact linguistics.

Zuckermann (2009) provides numerous instances of calquing in Modern Hebrew and analyses them in the light of the Congruence Principle.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	POSS	possessive pronoun
DEF	definite article	PRG	pragmatic marker
F	feminine	PRF	perfect
GEN	genitive	PRO	pronoun
M	masculine	PROX	proximal
NC	noun class	RL	recipient language
OBL	oblique	REFL	reflexive
PASS	passive	SG	singular
PAST	past	SL	source language
PL	plural	ABS	absolute state

References

- Alamin, Suzan. 2015. The Tagoi pronominal system. *Occasional Papers in the Study of Sudanese Languages* 11. 17–30.
- Aquilina, Joseph. 1987. *Maltese–english dictionary*. Malta: Midsea Books.
- Bauer, Laurie. 2001. Compounding. In Martin Haspelmath & Wolfgang Oesterreicher (eds.), *Language typology and universals*, 695–707. Berlin: De Gruyter.
- Borg, Albert & Marie Azzopardi-Alexander. 1997. *Maltese*. London: Routledge.
- Chikovani, Guram. 2005. Linguistic contacts in central asia. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: case studies from iranian, semitic and turkic*, 127–36. London: Routledge.
- Comrie, Bernard. 2011. Creoles and language typology. In Claire Lefebvre (ed.), *Creoles, their substrates and language typology*, 599–611. Amsterdam: John Benjamins.
- El Aissati, Abderahman. 2011. Berber loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of arabic language and linguistics*. Leiden: Brill.
- Haspelmath, Martin. 2009. Lexical borrowing: concepts and issues. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world’s languages: A comparative handbook*, 35–54. Berlin: De Gruyter.
- Heath, Jeffrey. 1984. Language contact and language change. *Annual Review of Anthropology* 13. 367–84.
- Heine, Bernd. 2011. Areas of grammaticalization and geographical typology. In Osamu Hieda, Christa König & Hiroshi Nakagawa (eds.), *Geographical typology and linguistic areas with special reference to africa*, 41–66. Amsterdam: John Benjamins.

Stefano Manfredi

- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Keesing, Roger. 1988. *Melanesian Pidgin and the Oceanic substrate*. Stanford, CA: Stanford University Press.
- König, Ekkehard, Peter Siemund & Stephan Töpper. 2013. Intensifiers and reflexive pronouns. In Matthew Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Kossmann, Maarten. 2000. *Esquisse grammaticale du rifain oriental*. Leuven: Peeters.
- Lefebvre, Claire. 1998. *Creole genesis and the acquisition of grammar*. Cambridge: Cambridge University Press.
- Manfredi, Stefano. 2010. *A grammatical description of kordofanian baggara Arabic*. Naples: Università degli Studi di Napoli 'L'Orientale' doctoral dissertation.
- Manfredi, Stefano. 2017. *Arabi ĵuba: Un pidgin-créole du Soudan du Sud*. Leuven: Peeters.
- Manfredi, Stefano. 2018. Arabic as a contact language. In Reem Bassiouney & Elabbas Benmamoun (eds.), *The Routledge handbook of Arabic linguistics*, 407–20. London: Routledge.
- Mettouchi, Amina. 2012. Kabyle corpus. corpus recorded, transcribed and annotated by a. mettouchi. In Amina Mettouchi & Christian Chanard (eds.), *The Cor-pAfroAs corpus of spoken AfroAsiatic languages*. doi.org/10.1075/scl.68.website.
- Meyerhoff, Miriam. 2009. Replication, transfer, and calquing: using variation as a tool in the study of language contact. *Language Variation and Change* 21. 297–317.
- Nakao, Shuichiro. 2012. Revising the substratal/adstratal influence on Arabic creoles. In Osamu Hieda (ed.), *Challenges in nilotic linguistics: phonology, morphology, and syntax*, 127–49. Tokyo: ILCAA.
- Owen, Roger. 1909. *Bari grammar and vocabulary*. London: Bumpus.
- Owens, Jonathan. 2014. Many heads are better than one: the spread of motivated opacity via contact. *Linguistics* 52(1). 125–165.
- Owens, Jonathan. 2015. Idioms, polysemy, context: a model based on Nigerian arabic. *Anthropological Linguistics* 57(1). 46–98.
- Owens, Jonathan. 2016. The lexical nature of idioms. *Language Sciences* 57. 49–69.
- Pepper, Stephen. Forthcoming. *The typology of binominal lexemes: topic maps, railways and the nature of associative thought*. Oslo: University of Oslo doctoral dissertation.

- Ratcliffe, Robert. 2005. Bukhara arabic: a metatypized dialect of arabic in central asia. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: case studies from iranian, semitic and turkic*, 141–51. London: Routledge.
- Ritt-Benmimoun, Veronika, Smaranda Grigore, Jocelyne Owens & Jonathan Owens. 2017. Three idioms, three dialects, one history: egyptian, nigerian and tunisian arabic. In Veronika Ritt-Benmimoun (ed.), *Tunisian and libyan arabic dialects: common trends – recent developments – diachronic aspects*, 43–84. Zaragoza: Prensas de la Universidad de Zaragoza.
- Ross, Malcolm. 2007. Calquing and metatypy. *Journal of Language Contact* 1(1). 116–143.
- Sasse, Hans-Jürgen. 1992. Language decay and contact-induced change: similarities and differences. In Matthias Brenzinger (ed.), *Language death: factual and theoretical explorations with special reference to east africa*, 59–80. Berlin: Mouton de Gruyter.
- Souag, Lameen. 2017. Clitic doubling and language contact in Arabic. *Zeitschrift für Arabische Linguistik* 66. 45–70.
- Taine-Cheikh, Catherine. 2008. Arabe(s) et berbère en contact: le cas mauritanien. In Mena Lafkioui & Vermondo Brugnatelli (eds.), *Berber in contact: Linguistic and sociolinguistic perspectives*, 113–38. Cologne: Rüdiger Köppe.
- Taine-Cheikh, Catherine. 2012. Arabe(s) et berbère en mauritanie. In Gunvor Mejdell & Lutz Edzard (eds.), *High vs. low and mixed varieties: Status, norms and functions across time and languages*, 88–108. Wiesbaden: Harrassowitz.
- Thomason, Sarah G. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah G. & Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Vanhove, Martine. 1993. *La langue maltaise: etudes syntaxiques d'un dialecte arabe «périphérique»*. Wiesbaden: Harrassowitz.
- Vanhove, Martine, Catherine Miller & Dominique Caubet. 2009. The grammaticalisation of modal auxiliaries in maltese and arabic vernaculars of the mediterranean area. In Björn Hansen & Ferdinand de Haan (eds.), *Modals in the languages of europe: a reference work*, 325–362. Berlin: Mouton de Gruyter.
- Wilmsen, David & Manfred Woidich. 2011. Egypt. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Winford, Donald. 2003. *An introduction to contact linguistics*. Malden: Blackwell.

Stefano Manfredi

- Winford, Donald. 2005. Contact-induced change: classification and processes. *Diachronica* 22(2). 373–427.
- Zuckermann, Ghil'ad. 2009. Hybridity versus revivability: multiple causation, forms and patterns. *Journal of Language Contact* VARIA 2. 40–67.

Chapter 18

Contact and the expression of negation

Christopher Lucas

SOAS University of London

This chapter presents an overview of developments in the expression of negation in Arabic and a number of its contact languages, focusing on clausal negation, with some remarks also on indefinites in the scope of negation. For most of the developments discussed in this chapter, it is not possible to say for certain that they are contact-induced. But evidence is presented which, cumulatively, points to widespread contact-induced change in this domain being the most plausible interpretation of the data.

1 Overview of concepts and terminology

1.1 Jespersen's cycle

In the past few decades there has been steadily intensifying interest among linguists in historical developments in the expression of negation, with particular attention given to the fact that these developments typically give the appearance of being cyclical in nature. We can date the beginning of this sustained interest to Dahl's (1979) typological survey of negation patterns in the world's languages, in which he coined the term JESPERSEN'S CYCLE¹ for what is by now the best-known set of developments in this domain: the replacement of an original negative morpheme with a newly grammaticalized alternative, after a period in which the two may co-occur, prototypically resulting in a word-order shift from preverbal to postverbal negation. The best-known examples of Jespersen's cycle

¹The name was chosen in recognition of the early identification of this phenomenon by the Danish linguist Otto Jespersen in a (1917) article, though others did identify the same set of changes earlier: Meillet (1912), for example, but also, significantly for the present work, Gardiner (1904), who observed a parallel set of changes in Coptic and Arabic as well as French (cf. van der Auwera 2009).



Christopher Lucas

(both supplied, among others, by Jespersen himself in his 1917 work) come from the history of English (1), and French (2).

(1) English (Jespersen 1917: 9)

a. Stage I – Old English

ic **ne** secge

1SG NEG say.PRS.1SG

‘I do not say.’

b. Stage II – Middle English

I **ne** seye **not**

1SG NEG say.PRS.1SG NEG

‘I do not say.’

c. Stage III – Early Modern English

I say **not**

(2) French (Jespersen 1917: 7)

a. Stage I – Old French

jeo **ne** di

1SG NEG say.PRS.1SG

‘I do not say.’

b. Stage II – contemporary written French

Je **ne** dis **pas**

1SG NEG say.PRS.1SG NEG

‘I do not say.’

c. Stage III – contemporary colloquial French

Je dis **pas**

1SG say.PRS.1SG NEG

‘I do not say.’

Jespersen’s cycle has since been the subject of intensive investigation, especially in the languages of Europe (e.g. Bernini & Ramat 1992; 1996; Willis et al. 2013), but also beyond (e.g. Lucas 2007; 2009; 2013; Lucas & Lash 2010; Devos & van der Auwera 2013; van der Auwera & Vossen 2015; 2016; 2017), with a picture emerging of a marked propensity for instances of Jespersen’s cycle to be areally distributed, as we will see below in the discussion of Jespersen’s cycle in Arabic and its contact languages (§2).

18 *Contact and the expression of negation*

While Jespersen's cycle is the best known, best studied, and perhaps cross-linguistically most frequently occurring set of changes in the expression of negation, two other important types of changes must also be mentioned here: Croft's cycle and changes to indefinites in the scope of negation.

1.2 Croft's cycle

In a typologically-oriented (1991) article, Croft reconstructs from synchronic descriptions of a range of languages a recurring set of cyclical changes in the expression of negation. Unlike Jespersen's cycle, in which the commonest sources of new negators are nominal elements expressing minimal quantities, such as 'step' or 'crumb', or generalizing pronouns like '(any)thing', CROFT'S CYCLE (named for Croft by Kahrel 1996), involves the evolution of new markers of negation developed from negative existential particles. Croft (1991: 6) distinguishes the following three types of languages:

Type A: the verbal negator is also used to negate existential predicates

Type B: there is a special negative existential predicate distinct from the verbal negator

Type C: there is a special negative existential predicate, and this form is also used to negate verbs

For Type A, Croft (1991: 7) cites the example of Syrian Arabic *mā fī* 'there is not' and *mā baʃref* 'I do not know' among others. For Type B he cites (1991: 9) cites the exam), among other examples, the contrast between the Amharic negative existential *yällām* (affirmative existential *allu*) and regular verbal negation *a(l)...-əm*. For Type C he cites (1991: 11–12) cites the exam) Manam (Oceanic) among other languages, giving the example in (3).

(3) Manam (Croft 1991: 11–12; Lichtenberk 1983: 385, 499)

a. Verbal negation

tágo u-lóŋo

NEG(EXS) 1SG.REAL-hear

'I did not hear.'

b. Negative existential predicate

anúa-lo tamóata tágo [*i-sóaʔi]

village-in person NEG.EXS [3SG.REAL-EXS]

'There is no one in the village.'

Christopher Lucas

A number of languages also exhibit variation between two of the types: $A \sim B$, $B \sim C$, and $C \sim A$. This indicates a cyclical development $A > B > C > A$, in which a special negative existential predicate arises in a language ($A > B$), comes to function also as a verbal negator ($B > C$), and is then felt to be the negator proper, requiring supplementation by a positive existential predicate in existential constructions ($C > A$).

While Croft's cycle is less common than Jespersen's cycle, and has not been shown to have occurred in its entirety in the recorded history of any language, we mention it here because recent work by Wilmsen (2014: 174–176; 2016), discussed below in (§2.1.2), argues for several instances of Croft's cycle in the history of Arabic.

1.3 Changes to indefinites in the scope of negation

The final major set of common changes to be dealt with here involve indefinite pronouns and quantifiers in the scope of negation. Here too cyclical patterns are commonplace, and these changes have been labelled “the argument cycle” (Ladusaw 1993) or “the quantifier cycle” (Willis 2011). What we find is that certain items, typically quantifiers such as ‘all’ or ‘one’ or generic nouns such as ‘person’ or ‘thing’, are liable to develop restrictions on the semantic contexts in which they can occur, namely what are referred to as either downward-entailing or non-veridical contexts (see Giannakidou 1998 for details and the distinction between the two). In essence, this means interrogative, conditional, and negative clauses, as well as the complements of comparative and superlative adjectives, but not ordinary affirmative declarative clauses. Items that are restricted to appearing in such contexts, such as English *ever* (consider the ungrammaticality of, e.g., **I've ever been to Japan*), are generally termed NEGATIVE POLARITY ITEMS. Often, however, we find negative polarity items whose appearance is restricted to a subset of these contexts, and much the most common restriction is to negative contexts only. Items with this narrower distribution, such as the English degree adverbial phrase *one bit*, are generally termed strong negative polarity items and those with the wider downward-entailing/non-veridical distribution may be termed weak negative polarity items in contrast.

A commonly recurring diachronic tendency of such items is that they become stronger over time. That is, an item goes from having no restrictions, to being a weak negative polarity item, to being a strong negative polarity item, to eventually being itself inherently negative. The best-known instance of this progression comes from French *personne* ‘nobody’ and *rien* ‘nothing’. These derive from the ordinary, unrestricted Latin generic nouns *persona* ‘person’ and *rem* ‘thing’ and

18 Contact and the expression of negation

still behaved as such in medieval French, as in (4).

- (4) Medieval French (Hansen 2013: 72; Buridant 2000: 610)

Et si vous dirai une rien.
and so 2PL say.FUT.1SG INDF.SG.F thing
'And so I'll tell you a thing.'

In later medieval French they grammaticalized as indefinite pronouns and began to acquire a weak negative polarity distribution, as in the interrogative example in (5).

- (5) C13th French (Hansen 2013: 72; Buridant 2000: 610)

As tu rien fet?
AUX.2SG 2SG anything do.PTCP.PST
'Have you done anything?'

In present-day French these items have become essentially inherently negative, as shown in (6). They can no longer appear in interrogative, conditional or main declarative clauses with an affirmative interpretation (Hansen 2013: 73), though an affirmative interpretation remains possible in comparative complements, albeit largely in frozen expressions, as in *rien au monde* 'anything in the world' in (7).

- (6) Contemporary French (Hansen 2013: 68)

Qui t' a vu? Personne!
who 2SG.OBJ AUX.3SG see.PTCP.PAST nobody
'Who saw you? Nobody!'

- (7) Contemporary French (Hansen 2013: 73)

J' aime le vin mieux que rien au monde.
1SG like.PRS DEF.SG.M wine better than anything in+DEF.SG.M world
'I like wine better than anything in the world'

Note that French *rien* 'nobody' and *personne* 'nothing', like their equivalents in many other Romance varieties (e.g. Italian *niente* and *nessuno*), are not straightforward negative quantifiers like English *nobody* and *nothing*, even disregarding their behaviour in contexts such as (7). This is because French, like many other languages but unlike standard English, standard German, classical Latin etc., exhibits NEGATIVE CONCORD. This refers to the fact that when two (or more) elements which express negation on their own co-occur in a clause, the result is

Christopher Lucas

not logical double negation (i.e. a positive) but a single logical negative, as illustrated in (8).

- (8) Contemporary French (Hansen 2013: 69)
 Personne n' a rien dit
 nobody NEG AUX.PRS.3SG nothing say.PTCP.PST
 'Nobody said anything.'

Items which have this unstable behaviour are distinguished from straightforward negative items by the term *N-WORD* (coined by Laka 1990; see also Giannakidou 2006). We will see in (§3) that these distinctions and terminology are helpful in understanding developments in varieties of Arabic and its contact languages that directly parallel those described above for French.

2 Developments in the expression of clausal negation

2.1 Arabic

2.1.1 Synchronic description

One of the most striking ways that a number of spoken Arabic varieties differ from Classical and Modern Standard Arabic is in the expression of negation. In Classical and Modern Standard Arabic, and in the majority of varieties spoken outside of North Africa, negation is exclusively preverbal, with the basic verbal negator in the spoken varieties being *mā*, as in the Damascus Arabic example in (9).

- (9) Damascus Arabic (Cowell 1964: 328)
 hayy masʔale mā bəddahhək
 DEM.F matter NEG laugh.CAUS.IMPF.IND.3SG.M
 'This is not a laughing matter.' (lit. 'does not cause laughter')

But in the varieties spoken across the whole of coastal North Africa and into the southwestern Levant, as well as in parts of the southern Arabian Peninsula (see Diem 2014; Lucas 2018 for more precise details), negation is bipartite, with preverbal *mā* joined by an enclitic *-š* which follows any direct or indirect pronominal object clitics, as in (10).

- (10) Cairo Arabic (advertising slogan)

(11) Palestinian Arabic (Seeger 2013: 147)
 badaħħinⁱ-š
 smoke.IMP.F.IND.1SG-NEG
 ‘I don’t smoke.’

There is near unanimous agreement among those who have considered the matter that the bipartite construction illustrated in (10) arose from the preverbal construction via grammaticalization, phonetic reduction, and cliticization of *šay?* ‘thing’, and that the purely postverbal construction in (11) in turn arose from the bipartite construction via omission of the original negator *mā*. As such, Lucas (2007; 2009; 2018) and Diem (2014), among many others, view this as a paradigmatic case of Jespersen’s cycle.

²Wilmsen (2014) also attempts to trace his etymology back further to the Proto-Semitic third-person pronouns. Apart from the implausability of the putative semantic shift from definite pronoun to indefinite determiner, this reconstruction is untenable on phonological grounds (see Al-Jallad 2015 for details).

Christopher Lucas

A specific element of Wilmsen's proposals that we need to consider in some detail here before we proceed is his suggestion that, while in his view we should not see the developments in Arabic as an instance of Jespersen's cycle, we can discern in them an instance of Croft's cycle. As we will see below, this suggestion involves a distortion or misunderstanding of both the Arabic data and the sorts of patterns that constitute genuine instances of Croft's cycle, but the proposal has some *prima facie* plausibility, because of the existence in some dialects of the south and east of the Arabian Peninsula of an existential predicate *šī/šē/šay*, as in (12).

- (12) Northern Omani Arabic (Eades 2009: 92)
 ḥmīr šē l-ḥmīr barra
 donkey.PL EXS DEF-donkey.PL outside
 'There were donkeys... the donkeys were outside.'

Note that a similar element *šī* [ʔi:], with the same existential function, is found in the Modern South Arabian languages (MSAL) of Yemen and Oman, as in (13), from Mehri of Yemen.

- (13) Mehri of Yemen (Watson 2011: 31)
 šī fšē
 EXS lunch
 'Is there any lunch?'

Though Wilmsen (2014: 126) seems to view Arabic *šī* and Modern South Arabian *šī* as cognates, it is more likely that the presence of this item in the one set of varieties is the result of transfer from the other (cf. Al-Jallad 2015). The direction of transfer is unclear, however. At first glance, the fact that *šī* as an affirmative existential is found in essentially all of the MSAL spoken on the Arabian Peninsula, which have a long history of intensive contact with Arabic, but not in Soqotri, spoken on the island of Soqatra, where contact with Arabic is more recent and less intensive (Simeone-Senelle 2003), would appear to suggest that this is an innovation within Arabic originally, which was then transferred to just those MSAL with which there was most contact. On the other hand, the precise situation in Soqotri is perhaps instructive. Here the affirmative existential predicate is a unique form *ino*, while the negative existential predicate is *bīši* (Simeone-Senelle 2011: 1108). It is conceivable that the latter is a borrowing from Arabic, since affirmative existentials in *b-* are widespread in the Arabic dialects of Yemen. But a negative existential predicate *bīši* or similar is completely unattested in the Yemeni data provided by Behnstedt (2016: 346–348). This suggests,

18 Contact and the expression of negation

therefore, that: (i) existential *šī* is an original feature of MSAL; (ii) Soqotri is an example of a Type B language in Croft's typology, having innovated a new affirmative existential predicate *ino*, such that there is a special negative existential predicate that is neither identical to the verbal negator, nor simply a combination of the verbal negator with the affirmative existential predicate; and (iii) *šī* as an existential predicate in Arabic dialects is the result of transfer of MSAL *šī*.

This scenario is supported by the distribution of existential *šī* within Arabic varieties: the only clear cases are in dialects of Yemen and Oman with a history of contact with MSAL, and dialects of the Gulf whose speakers are known to have migrated there from Yemen or Oman (such as Šiḥḥī, §2.4). In various places Wilmsen tries to make a case for existential uses of *šī* outside this region, but this appears to be the result of confusion on his part between *šī* as a *bona fide* existential predicate and the existential presupposition that will inevitably be associated with the use of *šī* as an indefinite determiner (see, e.g., Heim 1988 on the semantics of indefinite noun phrases). For example, Wilmsen (2014: 123) cites Caubet's (1993a: 123, 1993b: 280) Moroccan Arabic examples in (14) as evidence of an existential use of *šī* as far west as Morocco. But there is no justification for Wilmsen's contradicting Caubet's uncontroversial analysis of *šī* as an indefinite determiner here: there are no existential predicates in these examples – the existence of the referents of the indefinite noun phrases is presupposed, not asserted.

(14) Moroccan Arabic (Caubet 1993a: 123, Caubet 1993b: 280)

- a. *ši nās kayāklū-ha*
 INDF people eat.IMP.F.REAL.3PL-3SG.F
 'Some people eat it.'
- b. *ši nās kaybyēw əl-lbən*
 INDF people like.IMP.F.REAL.3PL DEF-milk
 'Some people like milk.'

Nevertheless, *šī* does function as an existential predicate in a few Arabic varieties. The question, then, is whether a negated form of this predicate participates in a version of Croft's cycle, as Wilmsen maintains.

For the vast majority of Arabic varieties the answer is a clear no: these varieties straightforwardly belong to Type A of Croft's typology. The verbal negator (*mā*, *mā...-š*, or *-š*) is also used to negate existential predicates, as illustrated in (15) for Cairo Arabic.

(15) Cairo Arabic, personal knowledge

Christopher Lucas

- a. ma ʕamaltⁱ-š ḥāga
 NEG do.PRF.1SG-NEG thing
 ‘I didn’t do anything.’
- b. ma fī-š ḥāga
 NEG EXS-NEG thing
 ‘There is nothing.’

Wilmsen (2014: 173–175) suggests that Type B and Type C constructions can also be found, however. For Type B (“there is a special negative existential predicate, distinct from the verbal negator”; Croft 1991: 6), he cites Ṣanʕānī *māšī* and Moroccan *māši*. Ṣanʕānī *māšī* is certainly a negative existential predicate. But there is nothing special about it – it is a paradigmatic Type A construction, with the negation of the existential predicate (*šī*) performed by the verbal negator (*mā*). Moroccan *māši*, on the other hand, is the negator for nominal predicates (equivalent to *muš/miš/mū* in dialects east of Morocco). It is not a negative existential predicate at all, and, as discussed above, the /ši/ component of this item does not function as an existential in Moroccan, unlike in Ṣanʕānī and other southern Arabian varieties. The existence of *māši* in Moroccan Arabic is thus irrelevant to the question of whether this constitutes a Type B variety.³ Moroccan is a Type A variety: the positive existential predicate is *kāyn* and it is negated with the ordinary Moroccan verbal negator *ma...-š* (Caubet 2011).

Wilmsen’s identification of Arabic varieties of Type C (“there is a special negative existential predicate, which is identical to the verbal negator”; Croft 1991: 6) depends on the idea that the Arabic predicate negator *māši/muš/miš/mū* is a negative existential predicate, which, as we have seen, it is not. If it were, it would be true that there are Arabic varieties that are optionally of Type C, since in Cairo Arabic, among other varieties, it is possible to negate verbs with *miš* instead of the usual *ma...-š*, as Mughazy (2003) and others have pointed out. But Cairo *miš* (and Moroccan *māši*) are not negative existential predicates, and there is no evidence to suggest they ever were. Moreover, since the Ṣanʕānī negative existential predicate *māšī* also does not seem to be able to function as a verbal negator, there is little apparent merit in Wilmsen’s (2014) attempt to recast the history of negation in Arabic as an instance of Croft’s cycle.⁴

³Van Gelderen (2018) argues that the definition of Croft’s cycle should be expanded to encompass cases in which new negators arise from the univerbation of verbal negators with copulas and auxiliaries, as well as existentials. Wilmsen’s (2014) presentation of Croft’s Cycle makes no mention of any predicates other than existentials participating in the cycle, however.

⁴This is not to deny, however, that some Arabic dialects show some incipient Type B tenden-

18 *Contact and the expression of negation*

2.1.3 Internal or external?

It is clear from the above discussion that there is no reason to doubt the majority view of the emergence of negative *-š* as an instance of Jespersen's cycle. What is less clear and more controversial is the question of whether language contact played a role in triggering these developments, or whether this was a purely internal phenomenon (cf. Diem 2014: 11–12). This is an issue about which it is impossible to be certain given our present state of knowledge. Lucas & Lash (2010) make the case that contact did play a triggering role, however, and also provide arguments against the widely held view that, in the words of Lass (1997: 209), “an endogenous explanation of a phenomenon is more parsimonious [than one invoking contact – CL], because endogenous change must occur in any case, whereas borrowing is never necessary” (cf. also Lucas 2009: 38–43). Aside from this generalized reluctance to invoke contact in explanations of linguistic change unless absolutely necessary, another factor that is likely operative in the preference for seeing the Arabic developments as a purely internal phenomenon is ignorance of the wider picture of negative developments in Arabic and its contact languages. It is scarcely an exaggeration to say that everywhere an Arabic variety with bipartite negation is spoken, there is (or was) a contact language that also has bipartite negation, and – just as importantly – wherever Arabic dialects have only a single marker of negation, the local contact languages do too. The picture is similar in Europe, Ethiopia (Lucas 2009), Vietnam (van der Auwera & Vossen 2015), and many other places besides. There can therefore be no doubt that negative constructions, and especially bipartite negation (and hence Jespersen's cycle more generally), are particularly prone to diffusing through languages in contact. In the following sections I will briefly survey apparent instances of transfer of bipartite or postverbal negation in Arabic and Coptic, Arabic and MSAL, Arabic and Kumzari, Arabic and Berber, and Arabic and Domari. For more details see Lucas (2007; 2009; 2013) and Lucas & Lash (2010).

cies of a different kind. For example, Behnstedt (2016: 347) cites the northern Yemeni dialects of Rās Maḥall as-Sūdeh, Ḥammām ‘Alī and Afk, as varieties in which different morphemes are used in positive and negative existentials, albeit the negative construction used in each case is identical to that used for ordinary verbal negation. In a different context, Stefano Manfredi (personal communication) points out that many urban speakers of Sudanese Arabic use the item *māfīs*, borrowed from Egyptian Arabic, as a negative existential, while ordinary verbal negation is performed with preverbal *mā* alone (without postverbal *-š*).

Christopher Lucas

2.2 Arabic and Coptic

Based on an examination of evidence from Judaeo-Arabic documents preserved in the Cairo Genizah, among other sources of evidence, Diem (2014) comes to the conclusion that the Arabic bipartite negative construction found across coastal North Africa originated in Egypt between the tenth and eleventh centuries. This chronology and point of origin conforms closely with the conclusions I have drawn on this point in my own work (Lucas 2007; 2009; Lucas & Lash 2010), except that I have argued that what triggered the development of bipartite negation in Egypt was contact with Coptic (the name for the Egyptian language from the first century CE onwards), which, at the relevant period, had a frequently-occurring bipartite construction *ən...ən*, as illustrated in (16).

- (16) Coptic (Lucas & Lash 2010: 389)
 en ti-na-tsabo-ou an e-amənte
 NEG 1SG-FUT-teach-3PL NEG on-hell
 ‘I will not teach them about hell.’

The argument made in Lucas & Lash (2010) is that native speakers of Coptic acquiring Arabic as a second language must have encountered sentences negated with preverbal *mā* only, but which also contained after the verb *šī/šāy*, functioning either as an argument ‘(any)thing’ or an adverb ‘at all’,⁵ and interpreted this as the second element of the bipartite negative construction that their first-language Coptic predisposed them to expect. If this is correct, then the initial transfer of bipartite negation from Coptic to Arabic in Egypt should be understood as an instance of imposition under source-language agentivity, in the terms of Van Coetsem (1988; 2000), while the presence of bipartite negation in the dialects spoken across the rest of coastal North Africa, and the southwestern Levant, should be understood as the result of contact between neighbouring dialects of Arabic.

2.3 Arabic and Modern South Arabian

Diem (2014: 73) – like Obler (1990: 148) and, following her, Lucas (2007: 416) – suggests that bipartite negation in the southern Arabian Peninsula must have spread there from Egypt. This is possible, but historical evidence of significant early migration flows in this direction is lacking. The alternative explanation offered by

⁵Diem (2014) makes the case that *šī/šāy* had already developed an adverbial use at a very early stage, and that it is this adverbial use that should be seen as the form that was reanalysed as a negator.

18 *Contact and the expression of negation*

Lucas & Lash (2010) is that bipartite negation in the Arabic dialects of this region is an independent parallel development, here triggered by contact with MSAL, all mainland varieties of which have a bipartite negative construction of their own (or once had – some, such as Ḥarsūsi, have largely progressed to stage III of Jespersen’s cycle and lost the original preverbal negator), as illustrated in (17) for Omani Mehri.

- (17) Mehri of Oman (Johnstone 1987: 23)

əl təhələz b-ey laʔ
 NEG nag.IMPF.2SG.M with-1SG NEG
 ‘Don’t nag me!’

If this is correct, then here too, exactly as with the Coptic–Arabic contact in the previous section, we must have had an instance of transfer under source-language agentivity, with MSAL-dominant acquirers of Arabic imposing a bipartite construction on their second-language Arabic by reanalysing *šī/šay* as a negator. The key point is that in all dialects in which *šī/šay* functioned as an indefinite pronoun or adverb ‘at all’, the potential was there for reanalysis as the second element in a bipartite negative construction. But aside from in the dialects of Egypt and the southern Arabian Peninsula (and latterly dialects adjacent to Egyptian) this reanalysis never took place. Why the reanalysis did take place in Egypt and the southern Peninsula can be understood as being the result of the catalysing effect of contact with languages which themselves had a bipartite negative construction.⁶

2.4 Arabic and Kumzari

Kumzari is an Iranian language with heavy influence from both Arabic and MSAL that has only recently been described in detail (see van der Wal Anonby forthcoming). It is spoken on the Musandam Peninsula of northern Oman, where its primary contact language of recent times has been the Šihhī variety of Arabic (see Bernabela 2011 for a sketch grammar), which is clearly of the originally southern Arabian type described by Holes (2016: 18–32).

Šihhī Arabic has no Jespersen stage-II (bipartite) negative construction, but it has both a typical eastern Arabic stage-I construction with *mā*, as in (18a), perhaps due to recent influence from other Gulf Arabic varieties, alongside a unique

⁶For further discussion of the details of these changes, including the issues of the semantics and positioning in the clause of the second negative element in each of the three languages, see Lucas & Lash (2010: 395–401).

Christopher Lucas

(for Arabic) stage-III postverbal construction with *-lu*, as in (18b). The latter construction is apparently a straightforward transfer of the postverbal negator *laʔ/ləʔ* of MSAL (17).

(18) Šiḥḥī Arabic (Bernabela 2011: 87)

- a. mā mšēt ḥaṣāb əl-yōm
NEG go.PRF.1SG Khasab DEF-day
‘I didn’t go to Khasab today.’
- b. yqōl-lu bass il-kilmātēn
say.IMPF.3SG.M-NEG only DEF-words.DU
‘He doesn’t just say the two words.’

The Kumzari negator is the typical Iranian (and Indo-Iranian) *na*. What is less typical is that *na* occurs postverbally in Kumzari, as shown in (19).

(19) Kumzari (van der Wal Anonby forthcoming: 211)

- mām-ō kōr bur na
mother-DEF blind become.3SG.REAL NEG
‘The mother didn’t become blind.’

It seems very likely that contact with Šiḥḥī Arabic has played a role in this shift to postverbal negation, though not enough is known about the historical sociolinguistics of these two speech communities to say with confidence which of the two languages the agents of this change were dominant in.

2.5 Arabic and Berber

Berber languages are spoken from the oasis of Siwa in western Egypt in the east, across to Morocco and as far south as Burkina Faso. The most southerly of the Berber varieties – Tashelhiyt, spoken in southern Morocco, Zenaga, spoken in Mauritania, and Tuareg, spoken in southern Algeria and Libya, Niger, Mali and Burkina Faso – have only preverbal negation, as illustrated by the Tuareg example in (20).

(20) Tuareg (Chaker 1996: 10)

- ur igle
NEG leave.PFV.3SG.M
‘He didn’t leave.’

18 Contact and the expression of negation

These languages have, until recently, either had little significant contact with Arabic, or otherwise only with varieties such as Ḥassāniyya that have only preverbal negation with *mā*. All other Berber varieties which are in contact with Arabic varieties with bipartite negation also themselves have bipartite negation, illustrated for Kabyle (Algeria) in (21), or, in a few cases, purely postverbal negation, as in Awjila (Libya), illustrated in (22). The one exception is Siwa, which negates with preverbal *lā* alone – clearly a borrowing from a variety of Arabic, though which variety is not clear (see Souag 2009 for further discussion).

- (21) Kabyle (Rabhi 1996: 25)
 ul ittaggad kra
 NEG fear.AOR.3SG.M NEG
 ‘He is not afraid.’
- (22) Awjila (Paradisi 1961: 82)
 akellim iššen-ka amakan
 servant know.PFV.3SG.M-NEG place
 ‘The servant didn’t know the place.’
- (23) Siwa (Souag 2009: 58)
 lā gā-nūsd-ak
 NEG FUT-come.1PL-DAT.2SG
 ‘We won’t come to you.’

Different Berber varieties have postverbal negators with a range of different forms, but in most cases they either derive from two apparently distinct Proto-Berber items **ḳāra* and *(h)arā(t) both meaning ‘thing’ (Kossmann 2013: 332), or are transparent loans of Arabic *šay/ši*. This fact, when combined with the respective geographical distributions of single preverbal and bipartite negation in Arabic and Berber varieties, is sufficient to conclude that the presence of bipartite negation in Berber is in large part a result of calquing the second element of the Arabic construction, pace Brugnatelli (1987) and Lafkioui (2013b) (see also Kossmann 2013: 334; and see Lucas 2007; 2009 for more detailed discussion).⁷ Given that, until recently, native speakers of Arabic in the Maghreb acquiring Berber as a second language will always have been greatly outnumbered by native speakers of Berber learning Arabic as a second language, we must assume that

⁷Another postverbal negator – Kabyle *ani* – derives from the word for ‘where’ (Rabhi 1992), and so should perhaps be seen as more of an internal development, or at least less directly contact-induced. Tarifit also has a postverbal negator *bu*, whose etymology is uncertain, but which has also been transferred to the Moroccan Arabic dialect of Oujda (Lafkioui 2013a).

Christopher Lucas

the agents of this change were Berber-dominant speakers who made the change under recipient-language agentivity in a process akin to what Heine & Kuteva (2005) call polysemy copying and contact-induced grammaticalization (see also, Leddy-Cecere, this volume; Manfredi, this volume; Souag, this volume).

2.6 Arabic and Domari

The final instance of contact-induced changes to predicate negation to be mentioned here concerns the Jerusalem variety of the Indo-Aryan language Domari, as described by Matras (1999; 2007; 2012; this volume).

Matras (2012: 350–351) describes two syntactic contexts in which negators borrowed from Palestinian Arabic are the only options in this variety of Domari. The first is with Arabic-derived modal auxiliaries that take Arabic suffix inflection, as in *bidd-* ‘want’ in (24). Here negation is typically with the Palestinian Arabic stage-III construction *-š* (without *mā*), as it is would be also in Palestinian Arabic.

- (24) Jerusalem Domari (Matras 2012: 351)
 ben-om bidd-hā-š žawwiz-hōš-ar
 sister-1SG want-3SG.F-NEG marry-VITR.SUBJ-3SG
 ‘My sister doesn’t want to marry.’

The second is when the negated predicate is nominal, as in (25a), or, to judge from Matras’s examples, when we have narrow focus of negation with ellipsis, as in (25b). Here the negator that would be used in these contexts in Arabic – *miš* – is transferred to Domari and functions in the same way.

- (25) Jerusalem Domari (Matras 2012: 350)
 a. bay-os mišš kury-a-m-ēk
 mother-3SG NEG house-OBL.F-LOC-PRED.SG
 ‘His wife is not at home.’
 b. day-om min ʕammān-a-ki mišš min ʕēl-oman-ki
 mother-1SG from Amman-OBL.F-ABL NEG from family-1PL-ABL
 day-om
 mother-1SG
 ‘My mother is from Amman, she’s not from our family, my mother.’

In addition to these straightforward borrowings, Domari has a bipartite negative construction in which both elements involve inherited lexical material, as illustrated in (26).

18 *Contact and the expression of negation*

(26) Jerusalem Domari (Matras 2012: 117)

ʕašān ihne ama n-mang-am-san-eʔ l-ʕarab

because thus 1SG NEG-want-1SG-3PL-NEG DEF-Arabs

‘Because of this I don’t like the Arabs.’

In Lucas (2013: 413–414) I pointed out that the second element of this construction – *-eʔ* – was apparently not attested in varieties of Domari spoken outside of Palestine, and suggested that its presence in Jerusalem Domari could therefore be the result of influence from the Palestinian bipartite negative construction. Herin (2016; 2018), however, has since convincingly shown that this is incorrect, and that the Jerusalem Domari bipartite construction is an internal development with cognates in more northerly varieties, the latter being in contact with Arabic varieties that lack the bipartite negative construction. What is unique about the Jerusalem variety of Domari is that here a stage-III construction with *-eʔ* alone is possible, omitting the original preverbal negator *n(a)* that appears in (25b). Herin (2018: 32) argues that it is this stage-III construction, not the stage-II bipartite construction, that should be seen as the result of contact with Palestinian Arabic.

Overall, therefore, while the details naturally vary from one contact scenario to another, we see that negative constructions appear just as liable to be transferred between varieties of Arabic and neighbouring languages as they are between the languages of Europe and beyond.

3 Developments in indefinites in the scope of negation

3.1 Loaned indefinites

The organization and behaviour of indefinites in the scope of negation seem to be much more resistant to transfer between languages than is the expression of clausal negation, at least in the case of Arabic and its contact languages.⁸ Direct borrowing of individual indefinite items is rather common, however. I make no attempt at an exhaustive list here, but note the following two examples for illustrative purposes.

First, Berber varieties stand out as frequent borrowers of Maghrebi Arabic indefinites. The negative polarity item *ħadd/ħədd* ‘anyone’ is borrowed by at least

⁸Though for recent discussion of a related case – namely the acquisition of a determiner function by the Berber indefinite *kra* ‘something, anything’ via a calque of the polyfunctionality of Maghrebi Arabic *ši* – see Souag (2018).

Christopher Lucas

Siwa (Souag 2009: 58), Kabyle, Shawiya, Mozabite (Rabhi 1996: 29), and Tashelhiyt (Boumalk 1996: 41). The n-word *walu* ‘nothing’ is borrowed by at least Tarifit (Lafkioui 1996: 54), Tashelhiyt, and Central Atlas Tamazight (Boumalk 1996: 41). *ħatta*, in its function as an n-word determiner, is borrowed by at least Tashelhiyt (Boumalk 1996: 41). *qāl*, in its function as a negative polarity adverb ‘at all’, is borrowed by at least Tarifit and Central Atlas Tamazight (Boumalk 1996: 42). And the negative polarity adverb *ʕumr ‘(n)ever’ (< ‘age, lifetime’) is borrowed by at least Kabyle, Mozabite (Rabhi 1996: 30), and Tarifit (Lafkioui 1996: 72). Why these items should have been so freely borrowed, when each of them, with the possible exception of *ħatta*, have direct native equivalents, is unclear. But it is perhaps to be connected with the high degree of expressivity typically associated with negative statements containing indefinites, which therefore creates a constant need for new and “extravagant” (in the sense of Haspelmath 2000) means of expressing these meanings.

Second, while Arabic itself seems to have been much more constrained in its borrowing of indefinites from other languages, we can here point at least to the n-word *hič* ‘nothing’, borrowed from Persian, which Holes (2001: 549) includes in his glossary of pre-oil era Bahraini Arabic, citing also Blanc (1964: 159) and Ing-ham (1973: 547) for its occurrence in Baghdad and Khuzestan Arabic respectively. It remains in use in the latter (cf. Leitner, this volume), but consultations with present-day speakers of Baghdad Arabic indicate that, in this variety at least, this item has since dropped out of use.

3.2 The indefinite system of Maltese

While most or perhaps all Arabic varieties have at least some items that qualify as n-words according to the definition in §1.3, it is only Maltese that has developed into a straightforward negative-concord language with a full series of n-word indefinites in largely complementary distribution with a separate series of indefinites that cannot appear in the scope of negation, as is the situation in French, described in §1.3. These two series are shown in Table 1, adapted from Haspelmath & Caruana (1996: 215).

All the lexical material that makes up the Maltese indefinite system illustrated in Table 1 is inherited from Arabic, but the neat paradigm of n-words for determiner, ‘thing’, ‘person’, ‘time’, and ‘place’ is much more typical of European Romance languages than of Arabic. The extent to which, for example, *xejn* ‘nothing’ (deriving from *šay?* ‘thing’)⁹ is felt by Maltese speakers to be inherently

⁹Contrary to what I suggested in Lucas (2009: 83–84), the final segment of this item should

18 Contact and the expression of negation

Table 1: Maltese indefinites

	n-words	non-n-words
Determiner	ebda	xi
Thing	xejn	xi ħaġa
Person	ħadd	xi ħadd
Time	qatt	xi darba
Place	imkien	xi mkien

negative, is shown by the existence of the denominal verb *xejjen* meaning ‘to nullify’, as illustrated in (27).¹⁰

- (27) Maltese (Lucas 2013: 441)
Iżda xejjen lil-u nnifs-u
but nullify.PRF.3SG.M OBJ-3SG.M self-3SG.M
‘But he made himself nothing.’

As such, it seems likely that the intensive contact that occurred over several centuries between Maltese and the negative-concord languages Sicilian and Italian (cf. Lucas and Čěplö, this volume) played a role in these developments in the Maltese indefinite system. Precisely how this influence was mediated is hard to say, since both borrowing under recipient-language agentivity and imposition under source-language agentivity were likely operative in the Maltese–Romance contact situation, and either are possible here. See Lucas (2013: 439–444) for further discussion.

4 Conclusion

As we have seen, the overall areal picture of bipartite clausal negation in Arabic and its contact languages (and also, to a lesser extent, indefinites in the scope of negation) strongly suggests a series of contact-induced changes, and not a series of purely internally-caused independent parallel developments. What is required

not be understood as a fossilized retention of the indefinite suffix (so-called nunation or *tanwīn*) of Classical Arabic (cf. Lucas & Spagnol 2019).

¹⁰This is despite the fact that it may also occur in interrogatives with non-negative meaning (cf. Camilleri & Sadler 2017). Compare the French n-word *rien*, which, as illustrated in (7), retains a non-negative interpretation in a restricted set of negative-polarity contexts.

Christopher Lucas

in future research on this topic, to the extent that textual and other historical evidence becomes available, is a detailed, case-by-case examination of the linguistic and sociolinguistic conditions under which these constructions emerged in the languages in question. Such investigations would serve to either substantiate or undermine the contact-based explanations for these changes advanced in the course of this chapter. Ideally, they would also allow to understand in more detail the mechanisms of bilingual language use and acquisition that give rise to changes of this sort.

Further reading

Chaker & Caubet (1996) is an edited volume providing a wealth of descriptive data on the expression of negation in a number of Berber and Maghrebi Arabic varieties.

Diem (2014) is a detailed study of the grammaticalization of Arabic *šay?* as a negator, with particular attention paid to early sources of textual evidence for this development.

Willis et al. (2013) and **Breitbarth et al. (forthcoming)** are two volumes of a work examining in detail the history of negation in the languages of Europe and the Mediterranean.

Acknowledgements

The research presented in this chapter was partly funded by a Leadership Fellows grant from the UK Arts and Humanities Research Council, whose support is hereby gratefully acknowledged. I am also very grateful to Stefano Manfredi, Lameen Souag and Bruno Herin for their comments on an earlier draft of the chapter. Responsibility for any failings that remain is mine alone.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	M	masculine
ABL	ablative	OBJ	object
AOR	aorist	MSAL	Modern South Arabian
AUX	auxiliary	OBL	oblique
CAUS	causative	PASS	passive
DAT	dative	PST	past
DEF	definite article	PL	plural
DEM	demonstrative	PTCP	participle
DU	dual	PRED	predicate
EXS	existential	PRF	perfect (suffix conjugation)
F	feminine	PRS	present
FUT	future	REAL	realis
IMPF	imperfect (prefix conjugation)	SG	singular
IND	indicative	VITR	intransitive marker
INDF	indefinite		

References

Al-Jallad, Ahmad. 2015. What’s a caron between friends? A review article of Wilmsen (2014), with a special focus on the etymology of modern Arabic šī. *Bibliotheca Orientalis* 72(1/2). 34–46.

Behnstedt, Peter. 2016. *Dialect atlas of North Yemen and adjacent areas*. Leiden: Brill. (Translated by Gwendolin Goldbloom).

Bernabela, Roy S. 2011. *A phonology and morphology sketch of Šihhi Arabic dialect of alĠēdih (Oman)*. Leiden University M.A.

Bernini, Giuliano & Paolo Ramat. 1992. *La frase negativa nelle lingue d’Europa*. Bologna: Il Mulino.

Bernini, Giuliano & Paolo Ramat. 1996. *Negative sentences in the languages of Europe*. Berlin: Mouton de Gruyter.

Blanc, Haim. 1964. *Communal dialects in baghdad*. Cambridge, MA: Harvard University Press.

Boumalk, Abdallah. 1996. La négation en berbère marocain. In Salem Chaker & Dominique Caubet (eds.), *La négation en berbère et en arabe maghrébin*, 35–48. Paris: L’Harmattan.

Christopher Lucas

- Breitbarth, Anne, Christopher Lucas & David Willis (eds.). Forthcoming. *The history of negation in the languages of Europe and the Mediterranean*. Vol. II: *Patterns and processes*. Oxford: Oxford University Press.
- Brugnatelli, Vermondo. 1987. La negazione discontinua in berbero e in arabo-magrebino. In Giuliano Bernini & Vermondo Brugnatelli (eds.), *Atti della 4a giornata di Studi Camito-semitici ed Indoeuropei*, 53–62. Milan: Unicopli.
- Buridant, Claude. 2000. *Grammaire nouvelle de l'ancien français*. Paris: Sedes.
- Camilleri, Maris & Louisa Sadler. 2017. Negative sensitive indefinites in maltese. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '17 conference*, 146–166. Stanford, CA: CSLI Publications.
- Caubet, Dominique. 1993a. *L'arabe marocain*. Vol. 1: *Phonologie et morphosyntaxe*. Leuven: Peeters.
- Caubet, Dominique. 1993b. *L'arabe marocain*. Vol. 2: *Syntaxe et catégories grammaticales*. Leuven: Peeters.
- Caubet, Dominique. 2011. Moroccan Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Chaker, Salem. 1996. Remarques préliminaires sur la négation en berbère. In Salem Chaker & Dominique Caubet (eds.), *La négation en berbère et en arabe maghrébin*, 9–22. Paris: L'Harmattan.
- Chaker, Salem & Dominique Caubet (eds.). 1996. *La négation en berbère et en arabe maghrébin*. Paris: L'Harmattan.
- Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington DC: Georgetown University Press.
- Croft, William. 1991. The evolution of negation. *Journal of Linguistics* 27(1). 1–27.
- Dahl, Östen. 1979. Typology of sentence negation. *Linguistics* 17. 79–106.
- Devos, Maud & Johan van der Auwera. 2013. Jespersen cycles in Bantu: Double and triple negation. *Journal of African Languages and Linguistics* 34. 205–274.
- Diem, Werner. 2014. *Negation in Arabic: A study in linguistic history*. Wiesbaden: Harrassowitz.
- Eades, Domenyk. 2009. The Arabic dialect of a Šawāwī community of northern Oman. In Enam Al-Wer & Rudolf De Jong (eds.), *Arabic dialectology in honour of Clive Holes on the occasion of his sixtieth birthday*, 77–98. Leiden: Brill.
- Gardiner, Alan H. 1904. The word *iwn*³. *Zeitschrift für Ägyptische Sprache und Altertumskunde* 41. 130–135.
- Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins.

18 *Contact and the expression of negation*

- Giannakidou, Anastasia. 2006. N-words and negative concord. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, vol. 3, 327–391. Oxford: Blackwell.
- Hansen, Maj-Britt Mosegaard. 2013. Negation in the history of French. In David Willis, Christopher Lucas & Anne Breitbarth (eds.), *The history of negation in the languages of Europe and the Mediterranean*, vol. 1: *Case studies*, 51–76. Oxford: Oxford University Press.
- Haspelmath, Martin. 2000. The relevance of extravagance: A reply to bart geurts. *Linguistics* 38. 789–798.
- Haspelmath, Martin & Josephine Caruana. 1996. Indefinite pronouns in maltese. *Rivista di Linguistica* 8. 213–227.
- Heim, Irene. 1988. *The semantics of definite and indefinite noun phrases*. New York: Garland Publishing.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Herin, Bruno. 2016. Elements of domari dialectology. *Mediterranean Language Review* 23. 33–73.
- Herin, Bruno. 2018. The Arabic component in domari. In Stefano Manfredi & Mauro Tosco (eds.), *Arabic in contact*, 19–36. Amsterdam: John Benjamins.
- Holes, Clive. 2001. *Dialect, culture, and society in eastern Arabia*. Vol. 1: *Glossary*. Leiden: Brill.
- Holes, Clive. 2016. *Dialect, culture, and society in eastern Arabia*. Vol. 3: *Phonology, morphology, syntax, style*. Leiden: Brill.
- Ingham, Bruce. 1973. Urban and rural arabic in Khūzistān. *Bulletin of the School of Oriental and African Studies* 36. 533–553.
- Jespersen, Otto. 1917. *Negation in English and other languages*. Copenhagen: A. F. Høst.
- Johnstone, T. M. 1987. *Mehri lexicon and English–Mehri word-list*. London: Routledge.
- Kahrel, Peter. 1996. *Aspects of negation*. Amsterdam: University of Amsterdam doctoral dissertation.
- Kossmann, Maarten. 2013. *The Arabic influence on Northern Berber*. Leiden: Brill.
- Ladusaw, William A. 1993. Negation, indefinites, and the Jespersen cycle. In Barbara A. Kaiser Joshua S. Guenter & Cheryl C. Zoll (eds.), *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society*, 437–46. Berkeley: Berkeley Linguistics Society.

Christopher Lucas

- Lafkioui, Mena. 1996. La négation en tarifit. In Salem Chaker & Dominique Caubet (eds.), *La négation en berbère et en arabe maghrébin*, 49–78. Paris: L'Harmattan.
- Lafkioui, Mena. 2013a. Negation, grammaticalization and language change in North Africa: The case of the negator NEG _____*bu. In Gabriele Iannàccaro Giorgio Francesco Arcodia Federica Da Milano & Paolo Zublena (eds.), *Tilelli: Scritti in onore di Vermondo Brugnatelli*, 113–130. Rome: Caissa Italia.
- Lafkioui, Mena. 2013b. Reinventing negation patterns in moroccan Arabic. In Mena Lafkioui (ed.), *African Arabic: Approaches to dialectology*, 51–93. Berlin: De Gruyter Mouton.
- Laka, Itziar. 1990. *Negation in syntax: on the nature of functional categories and projections*. Cambridge: Massachusetts Institute of Technology doctoral dissertation.
- Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.
- Lichtenberk, Frantisek. 1983. *A grammar of Manam*. Honolulu: University of Hawaii Press.
- Lucas, Christopher. 2007. Jespersen's cycle in Arabic and berber. *Transactions of the Philological Society* 105(3). 398–431.
- Lucas, Christopher. 2009. *The development of negation in arabic and afro-asiatic*. Cambridge: University of Cambridge doctoral dissertation.
- Lucas, Christopher. 2013. Negation in the history of Arabic and Afro-Asiatic. In David Willis, Christopher Lucas & Anne Breitbarth (eds.), *The development of negation in the languages of Europe and the Mediterranean*, vol. 1: *Case studies*, 399–452. Oxford: Oxford University Press.
- Lucas, Christopher. 2018. On wilmsen on the development of postverbal negation in dialectal arabic. *Zeitschrift für Arabische Linguistik* 67. 44–70.
- Lucas, Christopher & Elliott Lash. 2010. Contact as catalyst: The case for coptic influence in the development of Arabic negation. *Journal of Linguistics* 46. 379–413.
- Lucas, Christopher & Michael Spagnol. 2019. Word-final /n/ in maltese. Paper presented at the 7th International Conference on Maltese Linguistics, Jagiellonian University, Kraków, 10–11 July.
- Matras, Yaron. 1999. The state of present-day domari in jerusalem. *Mediterranean Language Review* 11. 1–58.
- Matras, Yaron. 2007. Grammatical borrowing in domari. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 151–164. Berlin: Mouton de Gruyter.

18 Contact and the expression of negation

- Matras, Yaron. 2012. *A grammar of Domari*. Berlin: De Gruyter Mouton.
- Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia* 12. 384–400.
- Mughazy, Mustafa. 2003. Metalinguistic negation and truth functions: The case of Egyptian Arabic. *Journal of Pragmatics* 35(8). 1143–1160.
- Obler, Lorraine. 1990. Reflexes of classical Arabic *šay'un* 'thing' in the modern dialects. In James Bellamy (ed.), *Studies in Near Eastern culture and history in memory of Ernest T. Abdel-Massih*, 132–152. Ann Arbor: University of Michigan.
- Paradisi, Umberto. 1961. Testi berberi di augila (cirenaica). *Annali dell'Istituto Orientale di Napoli* 10. 79–91.
- Pat-El, Na'ama. 2016. David wilmsen, *Arabic indefinites, interrogatives, and negators: A linguistic history of western dialects*. *Journal of Semitic Studies* 61(1). 292–295.
- Rabhi, Allaoua. 1992. Les particules de négation dans la kabylie de l'est. *Etudes et documents berbères* 9. 139–145.
- Rabhi, Allaoua. 1996. De la négation en berbère: Les données algériennes. In Salem Chaker & Dominique Caubet (eds.), *La négation en berbère et en arabe maghrébin*, 23–34. Paris: L'Harmattan.
- Seeger, Ulrich. 2013. *Der arabische Dialekt der Dörfer um Ramallah, Teil 3: Grammatik*. Wiesbaden: Harrassowitz.
- Simeone-Senelle, Marie-Claude. 2003. Soqotri dialectology and the evaluation of the language endangerment. In S. A. Ba-Angood, M. O. Ba-Saleem & M. O. Hussein (eds.), *Proceedings of the second international symposium on the socotra archipelago*, 63–73. Aden: Aden University Printing.
- Simeone-Senelle, Marie-Claude. 2011. Modern South arabian. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1073–1113. Berlin: De Gruyter Mouton.
- Souag, Lameen. 2009. Siwa and its significance for Arabic dialectology. *Zeitschrift für Arabische Linguistik* 51. 51–75.
- Souag, Lameen. 2016. Werner diem: *Negation in Arabic: A study in linguistic history*. *Linguistics* 54(1). 223–229.
- Souag, Lameen. 2018. Berber etymologies in Maltese. *International Journal of Arabic Linguistics* 4(1). 190–223.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.

Christopher Lucas

- van der Auwera, Johan. 2009. The Jespersen cycles. In Elly van Gelderen (ed.), *Cyclical change*, 35–71. Amsterdam: John Benjamins.
- van der Auwera, Johan & Frens Vossen. 2015. Negatives between chamic and Bahnaric. *Journal of the Southeast Asian Linguistics Society* 8. 24–38.
- van der Auwera, Johan & Frens Vossen. 2016. Jespersen cycles in Mayan, Quechuan and Maipurean languages. In Elly van Gelderen (ed.), *Cyclical change continued*, 189–218. Amsterdam: John Benjamins.
- van der Auwera, Johan & Frens Vossen. 2017. Kiranti double negation: A copula conjecture. *Linguistics of the Tibeto-Burman Area* 40(1). 40–58.
- van der Wal Anonby, Christina. Forthcoming. *A grammar of kumzari: a mixed Perso-Arabian language of Oman*. Berlin: De Gruyter Mouton.
- van Gelderen, Elly. 2018. The negative existential and other cycles: Jespersen, Givón, and the copula cycle. Unpublished manuscript, Arizona State University. <https://ling.auf.net/lingbuzz/004051>.
- Watson, Janet C. E. 2011. South Arabian and Yemeni dialects. *Salford Working Papers in Linguistics and Applied Linguistics* 1. 27–40.
- Willis, David. 2011. Negative polarity and the quantifier cycle: Comparative diachronic perspectives from European languages. In Pierre Larrivé & Richard Ingham (eds.), *The evolution of negation: Beyond the Jespersen cycle*, 285–323. Berlin: De Gruyter.
- Willis, David, Christopher Lucas & Anne Breitbarth (eds.). 2013. *The history of negation in the languages of Europe and the Mediterranean*. Vol. I: Case studies. Oxford: Oxford University Press.
- Wilmsen, David. 2013. The interrogative origin of the Arabic negator -š: Evidence from copular interrogation in Andalusí Arabic, Maltese, and modern spoken Egyptian and Moroccan Arabic. *Zeitschrift für Arabische Linguistik* 58. 5–31.
- Wilmsen, David. 2014. *Arabic indefinites, interrogatives and negators: A linguistic history of western dialects*. Oxford: Oxford University Press.
- Wilmsen, David. 2016. Another Croft cycle in Arabic: The *laysa* negative existential cycle. *Folia Orientalia* 53. 327–367.

Chapter 19

Contact and variation in Arabic intonation

Sam Hellmuth

University of York

Evidence is emerging of differences among Arabic dialects in their intonation patterns, along known parameters of variation in prosodic typology. Through a series of brief case studies, this chapter explores the hypothesis that variation in intonation in Arabic results from changes in the phonology of individual Arabic varieties, triggered by past (or present-day) speaker bilingualism. If correct, variation in intonation should reflect prosodic properties of the specific languages that a particular regional dialect has had contact with.

1 Introduction

1.1 Contact-induced variation in intonation

The hypothesis explored in this chapter is that observed synchronic variation in intonation across Arabic dialects is contact-induced. In this scenario, differences between dialects would result from changes in the intonational phonology of individual varieties triggered by speaker bilingualism in Arabic and one or more other languages (Lucas 2015), either in the past, or up to and including the present day. To achieve this, I outline a framework for analysis of variation in intonation (§1.2), summarise recent research on the effects of bilingualism on the intonational phonology of bilingual individuals and the languages they speak (§1.3), and sketch the types of language contact scenario which may be relevant for Arabic (§1.4). In §2 I present case studies of prosodic features which appear to be specific to a particular dialect, on current evidence at least, and discuss which of the potentially relevant contact languages might have served as the potential



Sam Hellmuth

source of the feature in question, considering also possible endogenous (internal) sources of the change. The chapter closes (§3) with suggestions for future research.

1.2 Cross-linguistic variation in intonation

Any attempt to delimit the nature and scope of variation in intonation depends on the model of intonational phonology adopted. The analyses explored in §2 below are framed in the Autosegmental-Metrical (AM) theory of intonation (Ladd 2008), and the parameters of intonational variation explored are thus influenced by this choice.

A basic debate in analysis of intonation is whether the primitives of the system are whole contours (defined over an intonational phrase), or some sub-component of those contours (Ladd 2008). In AM theory, intonation is modelled as interpolation of pitch between tonal targets; these tonal targets are the primitives of the system and are of two types: pitch accents are associated with the heads of metrical domains (e.g. stressed syllables), boundary tones are associated with the edges of metrical domains (e.g. prosodic phrases). In AM, tonal targets are transcribed using combinations of high (H) or low (L) targets, which reflect significant peaks and valleys, respectively, in the pitch contour of the utterance; association of these events to landmarks in the metrical structure is marked using “*” for pitch accents (associated with stressed syllables) and “%” for boundary tones (associated with the right edge of prosodic phrases of different sizes). A typical AM analysis yields an inventory of the pitch accents and boundary tones needed to model the contours in a corpus of speech data, supported by a description of the observed contours (Jun & Fletcher 2015).

Ladd’s (2008) taxonomy of possible parameters of cross-linguistic variation in intonation (based on Wells 1982) envisages four broad (inter-related) categories of variation: systemic (differences in the inventory of pitch accents or boundary tones); semantic (differences in the meaning or function associated with a particular contour, pitch accent or boundary tone); realisational (differences in the phonetic realisation of otherwise parallel pitch accents or boundary tones); and phonotactic (differences in the distribution of pitch accents and boundary tones, or in their association to metrical structure).

Comparison of AM analyses across a typologically distinct set of languages (Jun 2005; 2015) has highlighted systematic cross-linguistic variation of a systemic and/or phonotactic nature, in terms of prosodic phrasing (with relatively smaller or larger domains involved in structural organization of intonation patterns), the distribution of tonal events relative to prosodic constituents (marking

either the edges or the metrical heads of phrases or both), and the size and composition of the inventory of tonal events regularly observed (pitch accents and boundary tones). There is also a large body of research on cross-linguistic variation in the phonetic realisation of pitch accents, in particular on peak alignment (Atterer & Ladd 2004; Ladd 2006) and scaling (Ladd & Morton 1997), confirming the existence of realisational cross-linguistic variation. The most advanced work on semantic variation to date has been on Romance languages, facilitated by a concerted effort to develop descriptions of these languages' intonation patterns within a common annotation system (Frota & Prieto 2015).

For Arabic, evidence is emerging of variation along similar lines. Recent review articles have highlighted clear differences in the size and composition of the inventory of pitch accents and boundary tones across Arabic dialects (Chahal 2011; El Zarka 2017), and in the association of pragmatic meanings with contours (cf. the case study in §2.1). Initial evidence suggests a difference between Jordanian and Egyptian Arabic in the mapping of prosodic phrases to syntax (Hellmuth 2016) similar to that reported across Romance languages (D'Imperio et al. 2005). Recent research suggests that Moroccan Arabic is a non-head-marking language in contrast to other Arabic dialects which are head-marking (see §2.2), mirroring the cross-linguistic variation captured in Jun's (2005) typology, and among the head marking dialects, there appears to be variation in the density of distribution of pitch accents (Chahal & Hellmuth 2015; see §2.3).

1.3 Contact-induced variation in intonation

A growing body of research has explored contact-induced prosodic change in the speech of bilingual communities and individuals. The initial focus of most studies was on second-language (L2) learners' intonation patterns, or studies of individual bilinguals (Queen 2012), and early L2 studies focused on realisational effects of a speaker's L1 on their L2, and vice versa (Atterer & Ladd 2004; Mennen 2004). More recent studies reveal a complex array of prosodic effects, both in terms of the features involved in the change (taking in all four of Ladd's categories of possible variation), and also in the directionality of effects (L1 on L2, L2 on L1, or hybrid effects).

Bullock (2009) characterizes the general contact-induced language change literature (e.g. Weinreich et al. 1968; Thomason & Kaufman 1988) as having made the assumption that segmental effects would "precede" prosodic effects, thus predicting prosodic effects would be seen only in contexts of widespread or sustained community bilingualism. As Bullock notes, however, there is no logical structural reason why this should be the case; her own study of English-like pros-

Sam Hellmuth

odic patterns in heritage French speakers in Pennsylvania confirms an effect of the dominant language in the prosodic domain (specifically in the realisation of focus) in speakers who in other respects maintain French segmental patterns.

Another example of prosodic properties of a dominant language affecting prosodic realisation of a heritage or second language, is that of immersion Gaelic learners in Scotland (Nance 2015). Nance demonstrates a structural change in progress in Gaelic, from lexical pitch accent – still used by older English–Gaelic bilinguals – to a purely post-lexical system, used by younger bilinguals in immersion education who produce Gaelic with English-like intonation. Similar effects of the dominant language on the non-dominant language are reported for Spanish in contact with Quechua (O'Rourke 2004).

The reverse effect has also been found in a number of studies, however, where prosodic properties of a non-dominant or heritage language have an effect on the prosodic realisation of the dominant language, in the speech of an individual or of the whole community. Fagyal (2005) studied a group of bilingual French–Arabic adolescents in Paris; instead of a typical French phrase-final rise, these speakers produce a phrase-final rise–fall contour in declaratives, similar to the contour observed in Moroccan Arabic (MA) in parallel contexts. Simonet (2011) shows that the steep (“concave”) final fall in Majorcan Catalan declaratives is now widely observed in Majorcan Spanish, replacing the typical gradual (“convex”) fall in Majorcan Spanish, but that each individual bilingual’s usage closely mirrors their reported language dominance. Colantoni & Gurlekian (2004) observe patterns of peak alignment in pre-nuclear accents and pre-focal downstep in Buenos Aires Spanish which differ from neighbouring varieties of Spanish but resemble those in Italian, and ascribe them to high levels of Spanish–Italian bilingualism in the city in the late nineteenth and early twentieth centuries. In this last case, the period of community bilingualism which triggered the change is now long past, but the effect on the prosodic patterns of the dominant language (in this case, Spanish) persists.

Finally, Queen (2001; 2012) reports a case of “fusion”: Turkish–German bilinguals in Germany display phrase-final intonation contours which are never used by monolinguals in either language, but are found only in the speech of bilinguals, in a new variety of German known as “Türkisch-Deutsch”.

This emerging literature suggests that contact-induced prosodic change is a frequent phenomenon, arising in varied forms and across diverse contact situations. Bullock (2009) suggests that prosody and intonation are especially prone to change for three reasons. First, because the acoustic parameters involved – pitch, intensity and duration – are part of the linguistic encoding of all languages, albeit

19 *Contact and variation in Arabic intonation*

in different constellations, and are thus readily adapted. Second, because, perceptually, all languages make use of prosodic parameters to convey some aspects of utterance-level meaning, thus the mapping of form to meaning is also readily adapted. Third, and perhaps most persuasively, because the form–meaning mapping in intonation is generally not fixed, but displays considerable inter- and intra-speaker variation (Cangemi et al. 2015; 2016) as well as contextual variation (cf. Walker 2014), and it is pockets of structural ‘indeterminacy’ of this kind which are prone to change in bilingual grammars (Sorace 2004). Queen (2001: 57) also suggests that the intertwining of form and function in intonation makes it a fruitful sphere for investigation of contact-induced change, because “intonation is one of the few linguistic elements that comments simultaneously on grammar, context and culture”. Indeed, Simonet’s (2011) work shows that speakers are able to adopt the intonation of a contact language without actually being proficient in the source language. Finally, Matras (2007) argues that the separate nature of prosody, which is processed separately from segmental phonology, and can be interpreted independently of the propositional content of the utterance, renders prosody more “borrowable” than other aspects of the grammar.

In sum, there is strong evidence that intonation patterns are highly porous, being transferred between dominant and non-dominant languages in either direction; intonation is thus a fruitful area for investigation of contact-induced language change. Among the literature reviewed here, the paper by Colantoni & Gurlekian (2004) most closely resembles the type of work which is needed in future for Arabic; they investigate present-day intonational variation in closely related varieties, and provide evidence from historical migration patterns to support the claim that the present-day variation can be ascribed to an earlier period of widespread bilingualism, in a language which is a plausible source of the feature in question. In the next section we outline a similar line of investigation for Arabic.

1.4 **Contact-induced variation in Arabic**

The time depth of descriptions of intonation patterns is shallow, due to the lack of historical audio recordings, and a general tendency that traditional grammars do not include detailed descriptions of prosody. It is thus difficult to reliably determine when changes in intonation may have happened, and the range of languages to be considered as the source of any putative intonational change is rather broad.

One set of potential source languages is the substrate languages spoken in a particular region before the arrival of Arabic, for example Amazigh (Berber) in

Sam Hellmuth

North Africa, and Egyptian–Coptic in Egypt. We might also see influence from the external languages which these indigenous languages were in contact with prior to the arrival of Arabic, such as Greek and Latin, or of other external languages whose influence was felt throughout the Arab world in later periods, such as Persian and Ottoman. Other possible source languages are European languages spoken along the northern coast of the Mediterranean, since large areas of southern Europe were under Arab rule for extended periods (7th–15th centuries CE), and contact through sea-borne trade is likely to have continued after that time. Conversely, large areas of the Arab world were under direct or indirect European control also (19th–20th centuries CE), and the influence of these languages is still felt today. Finally, we might also consider the potential effects of contact with global languages such as English, and with the L1 languages of migrant workers and long-term displaced language communities.

The decision to treat observed present-day variation as the result of change does not entail assuming that any one variety of Arabic was the ancestor of all dialects. Instead, the approach here will be to identify prosodic features which are seen in one Arabic dialect (or group of dialects), but not (yet) documented in any other dialects, as the most likely cases of potential contact-induced change. In each case study we evaluate the hypothesis by looking for evidence of the same feature in the relevant contact languages for the dialect in question, with comparison to possible language-internal sources of the change.

2 Contact-induced variation in Arabic intonation

2.1 Tunisian Arabic question marking

Tunisian Arabic polar questions are typically associated with a salient rise–fall pitch contour at the end of the utterance: speakers from southeast Tunisia produce a complete rise–fall in which pitch rises over the stressed syllable of the last word in the utterance to a peak, then falls to low; in contrast, speakers from Tunis produce a rise–plateau contour, in which, after the peak, pitch falls slightly then levels out. These patterns are illustrated in Figure 1 (Bouchhioua et al. 2019). The rise–fall prosodic contour is frequently accompanied by a segmental question marker, in the form of a vowel added to the end of the last word in the utterance. The quality of the epenthesis vowel is influenced partly by vowels earlier in the word (in a form of vowel harmony) and partly by regional dialect, though these vowel quality patterns require further investigation.

The rise–fall yes/no-question contour in TA differs from the rise seen in yes/no

19 Contact and variation in Arabic intonation

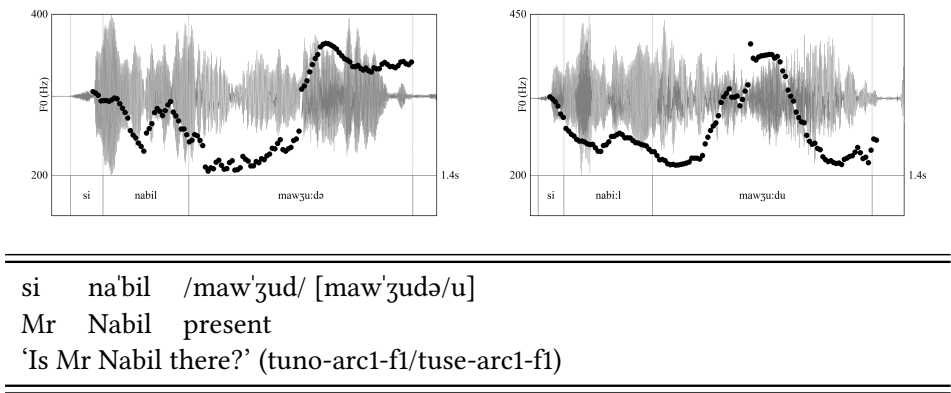


Figure 1: Pitch trace of yes/no questions from a Tunisian Arabic speaker from Tunis (left) and from the southeast (right).

questions in most Arabic dialects (Hellmuth to appear) and, in terms of distribution, from the rise–fall contour observed in Moroccan Arabic (MA) across all utterance types (not only in yes/no questions). The vowel epenthesis marker appears to be unique to TA, thus far.

A pattern of utterance-final vowel epenthesis has been observed in a number of Romance languages spoken along the northern edge of the Mediterranean, including Bari Italian (Grice et al. 2015), and different varieties of Portuguese (Frota et al. 2015). These cases of utterance-final vowel epenthesis are interpreted as text–tune adjustment, where segmental material is added to accommodate a complex prosodic contour. For example, in Standard European Portuguese, a more general rule of utterance-final vowel deletion is blocked in utterances bearing a complex prosodic contour, such as the fall–rise (H+L* LH%) on yes/no questions (Frota et al. 2015). In Bari Italian, epenthesis is seen on a range of utterance types, but – like Portuguese – its occurrence can be ascribed to tonal crowding (i.e. that the complex contour requires more segmental material to be realised). This is reflected in higher incidence of epenthesis on utterance-final monosyllables than on longer words, and on words in which the final sound is an obstruent than on words with a final sonorant (Grice et al. 2015).

Investigation of utterance-final vowel epenthesis in TA yes/no questions, in a corpus of data collected in Tunis, shows a very different pattern, however. In TA the incidence of epenthesis is not affected by the number of syllables in the utterance-final word nor by the type of final sound. In addition, whereas in the other Romance languages epenthesis occurs on a range of utterance types, in

Sam Hellmuth

TA epenthesis occurs only in yes/no-questions, and predominantly in yes/no-questions which are produced with a complex rise–plateau or rise–fall contour (Hellmuth forthcoming). The effects which in the Romance languages are taken as evidence of text–tune adjustment are lacking in TA, which appears to rule out a language-internal (endogenous) source of the TA pattern of vowel epenthesis.

The TA epenthetic vowel is in fact best characterized as an optional question marker comprising the vowel itself plus an accompanying fall in pitch. The segmental marker is well-known among Tunisian linguists, being described as “the pan-Tunisian question marker clitic [ā]” (Herin & Zammit 2017: 141), but the accompanying prosodic contour has received little attention in the literature until recently. This traditional question-marking strategy may however be in decline, since it now alternates with realisation of a yes/no-question using a simple rise contour similar to that found in most other Arabic dialects, and without an utterance-final epenthetic vowel. The picture is complicated by the fact that there is a somewhat higher incidence of epenthesis among young female speakers, who might be expected to use the traditional form less, rather than more (Hellmuth forthcoming).

The epenthesis + complex contour strategy in TA yes/no-questions stands out from other Arabic dialects and may thus be due to contact-induced prosodic change. Italian was spoken in Tunisia more widely than French, in the late nineteenth century (Sayahi 2011), and is thus a potential source of the contour, since rise–falls occur in yes/no questions in a number of Italian dialects (Gili Fivela et al. 2015). However, the conditioning environments of epenthesis reported for Bari Italian are very different, suggesting that contact with Italian is not a likely source of the epenthesis component of the TA pattern.

An alternative source of the vowel epenthesis pattern is French, since Tunisia has seen very high levels of bilingualism in TA and French from the late nineteenth century up to the present day, despite concerted efforts to reduce usage of French (Daoud 2007). Utterance-final schwa epenthesis has been reported as an emerging phenomenon in French (Hansen 1997), but its distribution is again much broader, being seen across a range of utterance types, and not restricted to yes/no questions. Despite clear evidence of contact-induced effects of French on TA in other domains, such as lexical borrowing, and a general trend towards use of French by female speakers (Walters 2011), the different distribution of final epenthesis in French suggests it is not the most likely source of the TA segmental question-marking strategy.

The other major contact language with TA is Tunisian Berber (TB). Although levels of TA–Berber bilingualism in Tunisia are now low, other than in certain

regions (Gabsi 2011), there was a sustained period of TA–TB bilingualism from the eleventh century, and TB is an important substrate of TA (Daoud 2007). Although there are no studies of the prosody of TB, to our knowledge, a recent detailed study of Zwara Berber, spoken close to the Tunisian border in western Libya, documents a polar question marking clitic /a/ which is obligatorily accompanied by a rise–fall contour (Gussenhoven 2017). The match of this description to the TA pattern is so close that it seems plausible that the TA question-marking pattern arose due to contact with TB during the period of sustained TA–TB bilingualism. The greater use of the epenthesis + contour strategy by female speakers than male speakers, as well as regional variation, makes this feature of TA ripe for further detailed sociolinguistic study.

2.2 Moroccan Arabic word prosody

Variation in word stress patterns across Arabic dialects has inspired much phonological investigation (Watson 2011), but the Moroccan Arabic (MA) stress system has defied analysis until recently. Mitchell (1993: 202) notes that “in contrast with all the other vernaculars [...], the place of prominence in a word in isolation is not carried over to its occurrence in the phrase and sentence”, and this characterization was confirmed experimentally by Boudlal (2001). A range of positions have emerged, with some authors claiming that MA does have word stress (Benkirane 1998; Burdin et al. 2014), and others that it does not (Maas 2013; El Zarka 2012).

It is now clear that MA is indeed typologically different from most other Arabic dialects in its word prosody. Whereas the majority of Arabic dialects have salient word-level stress and are thus clearly “head-marking” languages, in the typology proposed by Jun (2005), MA is a non-head-marking language in which tonal events mark the edges of prosodic phrases only. Bruggeman (2018) provides acoustic evidence that there are no consistent cues to lexical prominence in MA, and perceptual evidence that MA listeners display the same type of “deafness” to stress as has been reported for listeners in languages which also lack head marking such as French (Dupoux et al. 2001) and Persian (Rahmani et al. 2015).

Can this stark variation in prosodic type between MA and other dialects of Arabic be attributed to contact-induced change? The Arabic language has been in sustained contact with Amazigh (Moroccan Berber, MB) since the seventh century, but also with Latin, French and Spanish (Heath, this volume). Maas & Procházka (2012) argue from corpus data that MA and MB share a common phonology, across a range of segmental and suprasegmental features. Bruggeman (2018) confirms that there is no difference between MA and Tashlhiyt MB: both lack acoustic cues to word-level prominence in production and both groups

Sam Hellmuth

of listeners display stress deafness.

Since French is also an edge-marking language, without lexical stress, can we rule out French as an alternative source of this prosodic feature of MA? The main evidence comes from the fact that MA and MB also share other prosodic features which are not found in French, such as the shape of the tonal contour used to mark the edges of phrases, which is a rise in French (Delais-Roussarie et al. 2015), but a rise–fall in both Tashlhiyt MB (Grice et al. 2015; Bruggeman et al. 2017) and MA (Benkirane 1998; Hellmuth to appear). The contrast is also exemplified in Faygal’s (2005) study of French–MA bilinguals in Paris who use an MA rise–fall contour in French.

2.3 Egyptian Arabic accent distribution

Cairene Egyptian Arabic (EA) displays a rich distribution of sentence accents, with a pitch accent typically observed on every content word. This has been noted independently by different authors (Rifaat 1991; Rastegar-El Zarka 1997), and is observed in both read and spontaneous speech styles (Hellmuth 2006). Initial studies suggest that the same may be true also of some other dialects, such as Emirati (Blodgett et al. 2007) or Ḥiġāzi (Alzaidi 2014), but these observations await corroboration across different speech styles.

Dense accent distribution has been noted in some languages on the northern coast of the Mediterranean also, including Spanish and Greek (Jun 2005), although, in Spanish, the rich accent distribution seen in laboratory speech is reduced in spontaneous speech (Face 2003). Portuguese dialects vary in accent distribution: most varieties typically have an accent on every content word, but Standard European Portuguese shows an accent on the first and last words in an utterance only (Frota et al. 2015).

Rich accent distribution is not observed in Moroccan Arabic (Benkirane 1998), nor in Tunisian Arabic (Hellmuth to appear). If the EA accent distribution pattern were due to contact between EA and the southern European languages on the other side of the Mediterranean which share the tendency towards rich accent distribution, we might expect the pattern to be found all across North Africa.

There is strong documentary evidence from written sources of historical sustained multilingualism in Egypt. Greek arrived in Egypt in the fourth century BCE, serving as a formal administrative language alongside Egyptian for several centuries, and with the country reaching a state of “balanced societal bilingualism” in Greek and Egyptian in the sixth and seventh centuries CE (Papaconstantinou 2010: 6). Egyptian evolved into Coptic, and its prestige continued to increase from the sixth century CE onwards. After the Arab conquest in the seventh cen-

19 *Contact and variation in Arabic intonation*

tury CE, Arabic began to take over from Greek as the language of administration, eventually replacing Coptic in daily use (Papaconstantinou 2012).

Is it possible that Egyptian–Coptic or Greek is the source of the rich accent distribution observed in EA (and indeed in Romance languages in southern Europe)?

The distribution of full and long vowels in Coptic indicates that it had word-level prominence (Peust 1999: 270), but it is not possible to determine from written texts the nature or distribution of any tonal contours which may have been associated with prominent syllables. Anecdotal evidence suggests that the intonation patterns used in surviving liturgical forms of Coptic are very different from those in EA (Peust 1999: 32), though this difference may owe more to the liturgical setting than to properties of the languages in spoken form.

Ancient Greek is generally thought to have had a pitch accent system in which the primary marker of culminative accent in each word was pitch (Devine & Stephens 1985). The Koine Greek dialect used in Egypt is thought to have lost pitch accent in favour of a stress accent system, however, by the fourth century BCE (Benaissa 2012).

Support for the hypothesis that Greek is the original source of the rich accent distribution would come from a match between the historical spread of Koine Greek around the Mediterranean with the location of languages in which rich accent distribution is also found. This would predict that eastern varieties of Libyan (Cyrenaican) Arabic might also be found to display rich accent distribution. If rich accent distribution is confirmed in dialects of Arabic (such as Emirati or Ḥiǧāzi) which did not have sustained contact with Greek, or with EA more recently, this would argue against Greek as the original source. Although Nubi (Gussenhoven 2006) and Juba Arabic (Nakao 2013) display hybrid properties between stress and lexical tone, the most likely explanation of their prosodic patterns is direct contact with local tonal languages. A potential endogenous trigger for development of rich accent distribution would be the absence of other forms of phonological marking of word domains, which are indeed somewhat reduced in EA, in comparison to other dialects (Watson 2002), though the direction of causality of this correlation is not easily determined.

Accent distribution has only recently been added to the parameters of variation explored in work on prosody (Hellmuth 2007), and thus included in descriptions of the intonation systems of languages (e.g. Frota & Prieto 2015). As further descriptions emerge of more dialects of Arabic it will be important to include documentation of accent distribution, across genres and speaking styles, in future research.

Sam Hellmuth

3 Conclusion

There is much that we do not yet know about variation in intonation in Arabic, which leaves scope for investigation of further potential cases of contact-induced prosodic change. One such case may be the Syrian Arabic utterance-final rising intonation, sometimes known as “drawl”, which is found in yes/no questions but also across other utterance types (Cowell 1964), and which is an identifiable feature of the Damascus dialect (Kulk et al. 2003). Although the full geographical range of the pattern has not been investigated in detail, and may be diffused to other dialects in the Levant, this rising declarative intonation pattern stands out from most other Arabic dialects, and is thus another potential case of contact-induced change.

Another potential outlier pattern is the rise–fall intonation contour seen in yes/no questions in Yemeni Arabic from Ṣanʿāʾ (Hellmuth 2014). The full areal reach of this prosodic question-marking strategy is also not yet fully known, and may extend into Ḥiǧāzī Arabic and western dialects of Oman. However, we do know that a rise–fall is seen in both Tunisia and Morocco, though in these places the pattern may be due to contact with varieties of Amazigh Berber. Nevertheless, it is tempting to speculate how a pattern found in Yemen might also be found in Tunisia and Morocco, and thus to explore the potential role of contact-induced variation due to ancient migrations between the eighth and fourteenth centuries (Holes 2018).

Finally, the intonation of Modern Standard Arabic (MSA) and of other formal registers may prove to be a fruitful domain of future research. As our knowledge of the intonational phonology of spoken Arabic dialects improves, this will facilitate investigation of the extent to which the intonation patterns of a speaker’s regional dialect can be observed and/or perceived in their MSA speech, building on the findings of prior studies (El Zarka & Hellmuth 2009). An important goal would be to determine the extent to which a separate intonational system can be described for MSA, and to document the differential contribution to this system of specific genres of MSA discourse versus contact-induced influence due to widespread community mastery of multiple registers of the language.

All these investigations would benefit from improved documentation of the time depth of present-day surface intonation patterns. For the quasi-unique features explored in §2, we do not know whether these are the result of recent or much more distant historical change. This situation might be rectified through analysis of archive audio materials, though dialect studies have often worked on oral narratives, which yield only a limited range of prosodic expression (i.e.

usually few questions, and no information about turn-taking). A more viable strategy to gauge the time depth of contact-induced variation in Arabic intonation would be for future sociolinguistic studies to include prosodic features as variables in apparent time studies with participants in different age ranges, or for pre-existing corpora of apparent time data to be made available for prosodic analysis.

Further reading

There are two key reference works, so far, on intonation in Arabic dialects, based on secondary analysis of prior published work: [Chahal \(2011\)](#) and [El Zarka \(2017\)](#). [Hellmuth \(2019\)](#) suggests prosodic variables for inclusion in studies of variation and change in Arabic.

Acknowledgements

The Intonational Variation in Arabic corpus ([2017](#)) was funded by an award to the author by the UK Economic and Social Research Council (ES/I010106/1).

Abbreviations

AM	Autosegmental-Metrical theory	L1, L2	1st, 2nd language
		MA	Moroccan Arabic
BCE	before Common Era	MB	Moroccan Berber
CE	Common Era	MSA	Modern Standard Arabic
EA	Egyptian Arabic	TA	Tunisian Arabic
H	high tone	TB	Tunisian Berber
L	low tone	YA	Yemeni Arabic

References

Alzaidi, Muhammad. 2014. *Information structure and intonation in hijazi Arabic*. Colchester: University of Essex doctoral dissertation.

Atterer, Michaela & D. Robert Ladd. 2004. On the phonetics and phonology of “segmental anchoring” of f0: Evidence from German. *Journal of Phonetics* 32. 177–197.

Sam Hellmuth

- Benaissa, Amin. 2012. Greek language, education, and literary culture. In Christina Riggs (ed.), *The Oxford handbook of Roman Egypt*, 526–542. Oxford: Oxford University Press.
- Benkirane, Thami. 1998. Intonation in Western Arabic (Morocco). In Daniel Hirst & Albert Di Cristo (eds.), *Intonation systems: A survey of twenty languages*, 345–359. Cambridge: Cambridge University Press.
- Blodgett, Allison, Jonathan Owens & Trent Rockwood. 2007. An initial account of the intonation of emirati Arabic. In Jürgen Trouvain & William J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany*, 1137–1140. Saarbrücken: Universität des Saarlandes.
- Bouchhioua, Nadia, Sam Hellmuth & Rana Almbark. 2019. Variation in prosodic and segmental marking of yes–no questions in Tunisian Arabic. In Catherine Miller, Alexandrine Barontini, Marie-Aimée Germanos, Jairo Guerrero & Christophe Pereira (eds.), *Studies on Arabic dialectology and sociolinguistics. Proceedings of the 12th conference of AIDA held in Marseille from May 30th to June 2nd 2017*, 38–48. Aix-en-Provence: Institut de recherches et d'études sur le monde arabe et musulman.
- Boudlal, Abdelaziz. 2001. *Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic*. Rabat, Morocco: Université Mohammed V doctoral dissertation.
- Bruggeman, Anna. 2018. *Lexical and postlexical prominence in Tashlhiyt Berber and Moroccan Arabic*. Cologne: Universität zu Köln doctoral dissertation.
- Bruggeman, Anna, Timo B. Roettger & Martine Grice. 2017. Question word intonation in Tashlhiyt Berber: Is “high” good enough? *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1: 5). 1–27.
- Bullock, Barbara E. 2009. Prosody in contact in French: A case study from a heritage variety in the USA. *International Journal of Bilingualism* 13. 165–194.
- Burdin, Rachel Steindel, Sara Phillips-Bourass, Rory Turnbull, Murat Yasavul, Cynthia G. Clopper & Judith Tonhauser. 2014. Variation in the prosody of focus in head- and head/edge-prominence languages. *Lingua* 165. 254–276.
- Cangemi, Francesco, Martine Grice & Martina Krüger. 2015. Listener-specific perception of speaker-specific production in intonation. In Susanne Fuchs, Daniel Pape, Caterina Petrone & Pascal Perrier (eds.), *Individual differences in speech production and perception*, 123–145. Frankfurt Am Main: Peter Lang.
- Cangemi, Francesco, Dina El Zarka, Simon Wehrle, Stefan Baumann & Martine Grice. 2016. Speaker-specific intonational marking of narrow focus in Egyptian Arabic. In Jon Barnes, Alejna Brugos, Stefanie Shattuck-Hufnagel & Nanette

19 Contact and variation in Arabic intonation

- Veilleux (eds.), *Proceedings of Speech Prosody 2016, 31 May – 3 Jun 2016, Boston, USA*, 335–339. Boston: Boston University.
- Chahal, Dana. 2011. Intonation. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.
- Chahal, Dana & Sam Hellmuth. 2015. Comparing the intonational phonology of Lebanese and Egyptian Arabic. In Sun-Ah Jun (ed.), *Prosodic typology II: The phonology of intonation and phrasing*, 365–404. Oxford: Oxford University Press.
- Colantoni, Laura & Jorge Gurlekian. 2004. Convergence and intonation: Historical evidence from Buenos Aires Spanish. *Bilingualism: Language and Cognition* 7. 107–119.
- Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington DC: Georgetown University Press.
- D’Imperio, Mariapaola, Gorka Elordieta, Sónia Frota, Pilar Prieto & Marina Vigario. 2005. Intonational phrasing in romance: the role of syntactic and prosodic structure. In Sionia Frota, Marina Vigario & Maria João Freitas (eds.), *Prosodies, with special reference to Iberian languages*, 59–98. Berlin: De Gruyter.
- Daoud, Mohamed. 2007. The language situation in tunisia. In Robert B. Kaplan & Richard B. Jr. Baldauf (eds.), *Language planning and policy in Africa*, 256–307. Clevedon: Multilingual Matters.
- Delais-Roussarie, Elisabeth, Brechtje Post, Caroline Buthke M. Avanzi, Alberrrt Di Cristo, Ingo Feldhausen, Sun-Ah Jun, Philippe Martin, Trudel Meisenburg, Annie Rialland, Rafèu Sichel-Bazin & Hiyon Yoo. 2015. Intonational phonology of French: Developing a ToBI system for French. In Sónia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 63–100. Oxford: Oxford University Press.
- Devine, A. M. & Laurence D. Stephens. 1985. Stress in greek? *Transactions of the American Philological Association (1974–2014)* 115. 125–152.
- Dupoux, Emmanuel, Sharon Peperkamp & Núria Sebastián-Gallés. 2001. A robust method to study stress “deafness”. *Journal of the Acoustical Society of America* 110. 1606–1618.
- El Zarka, Dina. 2012. One *ra* – many meanings: syntax, semantics and prosody of the Moroccan modal particle *ra* and its Egyptian Arabic counterparts. *STUF – Language Typology and Universals. Sprachtypologie und Universalienforschung* 65(4). 412–431.
- El Zarka, Dina. 2017. Arabic intonation. In *Oxford Handbooks Online*. DOI: 10.1093/oxfordhb/9780199935345.013.77.
- El Zarka, Dina & Sam Hellmuth. 2009. Variation in the intonation of egyptian formal and colloquial arabic. *Langues et Linguistique* 22. 73–92.

Sam Hellmuth

- Face, Tim. 2003. Intonation in spanish declaratives: differences between lab speech and spontaneous speech. *Catalan Journal of Linguistics* 2. 115–131.
- Fagyal, Zsuzsanna. 2005. Prosodic consequences of being a beur: French in contact with immigrant languages in Paris. *University of Pennsylvania Working Papers in Linguistics* 10. 91–104.
- Frota, Sónia, Marisa Cruz, Flaviane Fernandes-Svartman, Gisela Collischonn, Aline Fonseca, Carolina Serra, Pedro Oliveira & Marina Vigário. 2015. International variation in Portuguese: European and Brazilian varieties. In Sónia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 235–283. Oxford: Oxford University Press.
- Frota, Sónia & Pilar Prieto (eds.). 2015. *Intonation in Romance*. Oxford: Oxford University Press.
- Gabsi, Zouhir. 2011. Attrition and maintenance of the berber language in Tunisia. *International Journal of the Sociology of Language* 211. 135–164.
- Gili Fivela, Barbara, Cinzia Avesani, Marco Barone, Giuliano Bocci, Claudia Crocco, Mariapaola D’Imperio, Rosa Giordano, Giovanna Marotta, Michelina Savino & Patrizia Sorianello. 2015. Intonational phonology of the regional varieties of Italian. In Sónia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 140–197. Oxford: Oxford University Press.
- Grice, Martine, Michelina Savino, Alessandro Caffo & Timo B. Roettger. 2015. The tune drives the text: schwa in consonant-final loan words in italian. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow, UK*. Glasgow: University of Glasgow. Paper number 0381.
- Gussenhoven, Carlos. 2006. Between stress and tone in Nubi word prosody. *Phonology* 23. 192–223.
- Gussenhoven, Carlos. 2017. Zwara (zuwārah) berber. *Journal of the International Phonetic Association* 48(3). 371–387.
- Hansen, Anita Berit. 1997. Le nouveau [ə] prépausal dans le français parlé à paris. In Jean Perrot (ed.), *Polyphonie pour Iván Fonágy*, 173–198. Paris: L’Harmattan.
- Hellmuth, Sam. Forthcoming. Text–tune alignment in tunisian arabic yes-no questions. In Marisa Cruz & Sónia Frota (eds.), *Prosodic variation (with)in languages: Intonation, phrasing and segments*. London: Equinox.
- Hellmuth, Sam. 2006. *Intonational pitch accent distribution in Egyptian Arabic*. London: SOAS doctoral dissertation.
- Hellmuth, Sam. 2007. The relationship between prosodic structure and pitch accent distribution: Evidence from egyptian arabic. *The Linguistic Review* 24. 289–314.

19 *Contact and variation in Arabic intonation*

- Hellmuth, Sam. 2014. Dialectal variation in Arabic intonation: Motivations for a multi-level corpus approach. In Samira Farwanah & Hamid Ouali (eds.), *Perspectives on arabic linguistics XXIV–XXV: papers from the annual symposia on Arabic linguistics, Texas, 2010 and Arizona, 2011*, 63–89. Amsterdam: John Benjamins.
- Hellmuth, Sam. 2016. Explorations at the syntax–phonology interface in arabic. In Stuart Davis & Usama Soltan (eds.), *Perspectives on arabic linguistics XXVII: papers from the Annual Symposium on Arabic Linguistics, Bloomington, Indiana, 2013*, 75–97. Amsterdam: John Benjamins.
- Hellmuth, Sam. 2019. Prosodic variation in arabic. In Enam Al-Wer & Uri Horesh (eds.), *The routledge handbook of arabic sociolinguistics*, 169–184. London: Routledge.
- Hellmuth, Sam. to appear. *Intonation in spoken Arabic dialects*. Oxford: Oxford University Press.
- Hellmuth, Sam & Rana Almbark. 2017. *Intonational Variation in Arabic Corpus*. UK Data Archive <http://reshare.ukdataservice.ac.uk/852878/>.
- Herin, Bruno & Martin R. Zammit. 2017. Three for the price of one: the dialects of kerkennah (tunisia). In Veronika Ritt-Benmimoun (ed.), *Tunisian and libyan arabic dialects: common trends – recent developments – diachronic aspects*, 135–146. Zaragoza: Prensas de la Universidad de Zaragoza.
- Holes, Clive. 2018. Introduction. In Clive Holes (ed.), *Arabic historical dialectology: linguistic and sociolinguistic approaches*, 1–28. Oxford: Oxford University Press.
- Jun, Sun-Ah. 2005. Prosodic typology. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 430–458. Oxford: Oxford University Press.
- Jun, Sun-Ah. 2015. Prosodic typology: by prominence type, word prosody, and macro-rhythm. In Sun-Ah Jun (ed.), *Prosodic typology II: The phonology of intonation and phrasing*, 520–539. Oxford: Oxford University Press.
- Jun, Sun-Ah & Janet Fletcher. 2015. Methodology of studying intonation: from data collection to data analysis. In Sun-Ah Jun (ed.), *Prosodic typology II: The phonology of intonation and phrasing*, 493–519. Oxford: Oxford University Press.
- Kulk, Friso, Cecilia Odé & Manfred Woidich. 2003. The intonation of colloquial Damascene Arabic: A pilot study. *Institute of Phonetic Sciences, University of Amsterdam, Proceedings* 25. 15–20.
- Ladd, D. Robert. 2006. Segmental anchoring of pitch movements: autosegmental phonology or gestural coordination? 18. 19–38.

Sam Hellmuth

- Ladd, D. Robert. 2008. *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladd, D. Robert & Rachel Morton. 1997. The perception of intonational emphasis: continuous or categorical? *Journal of Phonetics* 25. 313–342.
- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Maas, Utz. 2013. Die marokkanische akzentuierung. In Renaud Kutý, Ulrich Seeger & Shabo Talay (eds.), *Nicht nur mit engelszungen: beiträge zur semitischen dialektologie. festschrift für werner arnold zum 60. geburtstag*, 223–234. Wiesbaden: Harrassowitz.
- Maas, Utz & Stephan Procházka. 2012. Moroccan Arabic in its wider linguistic and social contexts. *STUF – Language Typology and Universals. Sprachtypologie und Universalienforschung* 65. 329–357.
- Matras, Yaron. 2007. The borrowability of structural categories. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 31–73. Berlin: Mouton de Gruyter.
- Mennen, Ineke. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32. 543–563.
- Mitchell, T. F. 1993. *Pronouncing Arabic 2*. Oxford: Clarendon Press.
- Nakao, Shuichiro. 2013. The prosody of Juba Arabic: Split prosody, morphophonology and slang. In Mena Lafkioui (ed.), *African Arabic: Approaches to dialectology*, 95–120. Berlin: De Gruyter Mouton.
- Nance, Claire. 2015. Intonational variation and change in Scottish Gaelic. *Lingua* 160. 1–19.
- O'Rourke, Erin. 2004. Peak placement in two regional varieties of Peruvian Spanish intonation. In Julie Auger, J. Clancy Clements & Barbara Vance (eds.), *Contemporary approaches to Romance linguistics: Selected papers from the 33rd Linguistic Symposium on Romance Languages (LSRL)*, 321–342. Amsterdam: John Benjamins.
- Papaconstantinou, Arietta. 2010. Introduction. In Arietta Papaconstantinou (ed.), *The multilingual experience in Egypt, from the Ptolemies to the Abbasids*, 1–16. London: Ashgate Publishing.
- Papaconstantinou, Arietta. 2012. Why did Coptic fail where Aramaic succeeded? Linguistic developments in Egypt and the Near East after the Arab conquest. In Alex Mullen & Patrick James (eds.), *Multilingualism in the Graeco-Roman Worlds*, 58–76. Cambridge: Cambridge University Press.

19 *Contact and variation in Arabic intonation*

- Peust, Carsten. 1999. *Egyptian phonology: An introduction to the phonology of a dead language*. Göttingen: Peust und Gutschmidt.
- Queen, Robin M. 2001. Bilingual intonation patterns: Evidence of language change from Turkish–German bilingual children. *Language in Society* 30. 55–80.
- Queen, Robin M. 2012. Turkish–German bilinguals and their intonation: Triangulating evidence about contact-induced language change. *Language* 88. 791–816.
- Rahmani, Hamed, Toni Rietveld & Carlos Gussenhoven. 2015. Stress “deafness” reveals absence of lexical marking of stress or tone in the adult grammar. *PLOS One* 10(12: e0143968).
- Rastegar-El Zarka, Dina. 1997. *Prosodische Phonologie des Arabischen*. Graz: Karl-Franzens-Universität doctoral dissertation.
- Rifaat, Khalid. 1991. *The intonation of arabic: an experimental study*. Alexandria: University of Alexandria doctoral dissertation.
- Sayahi, Lotfi. 2011. Introduction: current perspectives on tunisian sociolinguistics. *International Journal of the Sociology of Language* 211. 1–8.
- Simonet, Miquel. 2011. Intonational convergence in language contact: Utterance-final f0 contours in Catalan–Spanish early bilinguals. *Journal of the International Phonetic Association* 41. 157–184.
- Sorace, Antonella. 2004. Native language attrition and developmental instability at the syntax–discourse interface: Data, interpretations and methods. *Bilingualism: Language and Cognition* 7. 143–145.
- Thomason, Sarah G. & Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Walker, Traci. 2014. Form (does not equal) function: The independence of prosody and action. *Research on Language and Social Interaction* 47. 1–16.
- Walters, Keith. 2011. Gendering French in Tunisia: Language ideologies and nationalism. *International Journal of the Sociology of Language* 211. 83–111.
- Watson, Janet C. E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.
- Watson, Janet C. E. 2011. Word stress in arabic. In Marc Van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 5, 2990–3018. Oxford: Blackwell.
- Weinreich, Uriel, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfred Lehman & Yakov Malkiel (eds.), *Directions for historical linguistics*, 95–195. Austin: University of Texas Press.
- Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Name index

- Abd El-Jawad, Hassan, 309
 Abdel-Massih, Ernest, 324, 325
 Abdu, Hussein Ramadan, 76
 Abu Haidar, Farida, 190
 Abu-Haidar, Farida, 309
 Abuamsha, Duaa, 336
 Akın, Cahit, 40
 Akkuş, Faruk, 26, 30, 31, 34–38, 40, 117
 al-Ḥimyarī, 96
 Al-Essa, Aziza, 310, 311
 al-Idrīsī, Muḥammad, 233
 Al-Jallad, Ahmad, 68, 367, 368
 Al-Mahri, Abdullah Musallam, 257, 260
 Al-Qahtani, Khairiah, 311
 Al-Salman, Ibrahim, 175, 177
 Al-Shareef, Jamal, 309
 Al-Tamimi, Jalal, 314
 Al-Wer, Enam, 305, 309–314
 Alamin, Suzan, 352
 Albirini, Abdulkafi, 189, 193–198
 Alghamdi, Najla Manie, 310, 311, 314
 Almoaily, Mohammad, 177
 Alzaidi, Muhammad, 396
 Andersen, Henning, 322
 Anonymous, 291
 Appleyard, David, 217
 Aquilina, Joseph, 345
 Arvaniti, Amalia, 50
 Atlamaz, Ümit, 40
 Atterer, Michaela, 389
 Avram, Andrei, 101, 192, 348
 Avram, Andrei A., 102, 163, 165, 167, 173–181
 Azzopardi-Alexander, Marie, 102, 103, 112, 114–116, 192, 352
 Baider, Fabienne, 48
 Bakir, Murtadha J., 177, 181
 Bakker, Peter, 164, 175
 Ballou, Maturin M., 98
 Barbarossa, 290
 Aruj, 290
 Hizir, 290
 Bauer, Laurie, 344
 Beene, Wayne, 13
 Beeston, Alfred F. L., 246
 Behnstedt, Peter, 16, 65, 124, 146, 321, 331
 Belnap, R. Kirk, 197
 Ben Cheneb, Mohammed, 74
 Benaissa, Amin, 397
 Benkato, Adam, 84, 85, 97, 240
 Benkirane, Thami, 395, 396
 Benmamoun, Elabbas, 34, 35, 38, 40, 193–196
 Benoliel, José, 86
 Bergman, Elizabeth, 217
 Bergsträßer, Gotthelf, 313
 Bernabela, Roy S., 373, 374
 Bernini, Giuliano, 362
 Bettega, Simone, 223
 Biosca, Carles, 122

Name Name index

- Bisang, Walter, 326
 Biṭună, Gabriel, 26, 29, 30, 38
 Bizri, Fida, 175–179
 Blachère, Régis, 216
 Blanc, Haim, 23, 24, 37, 309
 Blau, Joshua, 112
 Blažek, Václav, 211, 222
 Blodgett, Allison, 396
 Blouet, Brian, 98
 Boogert, Nico van den, 231, 233, 240
 Borg, Albert, 102, 103, 112, 114–116, 118, 119, 192, 352
 Borg, Alexander, 48, 51, 54, 55, 57, 58, 61, 97, 99, 113, 114, 117, 124, 192, 333
 Boucherit, Aziza, 330
 Bouchhioua, Nadia, 392
 Boudelaa, Sami, 193
 Boudlal, Abdelaziz, 395
 Boumalk, Abdallah, 378
 Boumans, Louis, 188–190, 194, 197
 Bovington, R., 118, 119
 Brahimi, Fadila, 240
 Breitbarth, Anne, 380
 Brincat, Joseph M., 96, 97, 107, 118
 Broß, Michael, 137
 Broughton, Elizabeth, 297–299
 Browne, Gerald M., 207
 Bruggeman, Anna, 395, 396
 Brugnatelli, Vermondo, 240, 375
 Brunot, Louis, 86
 Brustad, Kristen, 10, 39, 196
 Bulakh, Maria, 257
 Bullock, Barbara E., 389, 390
 Burdin, Rachel Steindel, 395
 Bureng Vincent, George, 169
 Buridant, Claude, 365
 Bybee, Joan, 324, 329, 330, 332, 333
 Cadora, F. J., 308
 Camilleri, Maris, 115, 117, 379
 Camus Bergareche, Bruno, 294
 Canagarajah, Suresh A., 187
 Canavan, Alexandra, 146
 Cangemi, Francesco, 391
 Castellanos, Carles, 122
 Caubet, Dominique, 96, 190, 330, 369, 370, 380
 Čéplö, Slavomír, 28, 37, 55, 58, 65, 89, 91, 112, 115, 126, 192³, 345, 352, 379
 Chahal, Dana, 389, 399
 Chaker, Salem, 374, 380
 Chetrit, Joseph, 86
 Chikovani, Guram, 355
 CIA, 232
 Cifoletti, Guido, 74, 287, 291, 294
 Cinque, Guglielmo, 115
 Claudu, Ulrike, 324
 Clissold, Stephen, 290
 Coghill, Eleanor, 27, 114
 Cohen, David, 212, 213, 222
 Colantoni, Laura, 390, 391
 Colin, Georges S., 74
 Colin, Georges Séraphin, 123
 Comrie, Bernard, 118–120, 125, 354
 Cook, Albert R., 165
 Corré, Alan D., 299
 Corriente, Federico, 90
 Cotter, William M., 66, 309, 313
 Council of Europe, 49
 Cowell, Mark W., 366, 398
 Croft, William, 363, 370
 D’Anna, Luca, 189, 190, 192, 193, 195, 197
 D’Imperio, Mariapaola, 389
 Daher, Nazih Y., 190

Name Name index

- Dahl, Östen, 321, 326
 Dakhliia, Jocelyne, 291
 Dalli, A., 118, 119
 Dan, Pierre, 290
 Daniel, E. L., 3
 Daoud, Mohamed, 394, 395
 Davey, Richard J., 256, 331
 Davies, Humphrey, 324, 325
 Davis, Robert C., 288
 De Angelis, Pietro, 171
 de Cat, Cécile, 116
 De Planhol, X., 4
 de Prémare, André-Louis, 90
 de Ruiter, Jan Jaap, 188–190, 194, 195
 de Sacy, Silvestre, 96
 de Soldanis, Giovanni Pietro Francesco
 Agius, 95
 Delais-Roussarie, Elisabeth, 396
 Dendien, Jacques, 297
 Devine, A. M., 397
 Devos, Maud, 362
 Diem, Werner, 366, 367, 371, 372, 380
 Diessel, Holger, 147
 Dimitrakou, Dimitriou, 123
 Dodsworth, Robin, 151
 Döhla, Hans-Jörg, 114
 Dozy, Reinhart, 90
 Drewes, A. J., 125
 Drop, Hanke, 146
 Drower, E.S., 124
 Dupoux, Emmanuel, 395
 D’Anna, Luca, 75, 76
 D’Aranda, Emanuel, 290, 291
 D’Arvieux, Laurent, 290, 291

 Eades, Domenyk, 368
 Edzard, Lutz, 246
 Ehala, Martin, 188
 El Aissati, Abderahman, 350

 El Arifi, Samir, 230
 El Zarka, Dina, 389, 395, 398, 399
 El-Sayed, Rafed, 207
 Elias, Edward, 324, 325
 Elias, Elias, 324, 325
 Epps, Patience, 328
 Erwin, Wallace M., 14
 Extra, Guus, 189

 Fabri, Ray, 107, 115
 Face, Tim, 396
 Fagyal, Zsuzsanna, 390
 Fenech, Edward, 118
 Ferguson, Charles A., 123, 306
 Fiorini, Stanley, 98
 Flege, James E., 103
 Fletcher, Janet, 388
 Frajzyngier, Zygmunt, 148
 Frank, Louis, 292, 298, 299
 Friedman, Victor A., 116
 Frota, Sónia, 389, 393, 396, 397

 Gabriel, Christoph, 104
 Gabsi, Zouhir, 395
 Gafter, Roey, 305
 Garcès, Maria Antonia, 298
 Gardani, Francesco, 105
 Gardiner, Alan H., 361
 Gardner-Chloros, Penelope, 298
 Gasparini, Fabio, 223, 251, 254
 Gaszi, Dénes, 3
 Gatt, Albert, 96, 107
 Gaudefroy-Demombynes, Maurice, 216
 Gautier, E.-F, 230
 Gazsi, Dénes, 3–5
 Geary, Jonathan, 106
 Gendrot, Cédric, 250
 Gesenius, Wilhelm, 96
 Giancarli, Gigio Artemio, 289

Name Name index

- Giannakidou, Anastasia, 364, 366
 Gibson, Maik, 66, 96, 103, 104
 Giles, Howard, 188
 Gili Fivela, Barbara, 394
 Gilson, Erika H., 295
 Gonzo, Susan, 190, 191, 197, 200
 Goodwin, Stefan, 98
 Grandchamp, Pierre, 296
 Grech, Paul, 96
 Grice, Martine, 104, 393, 396
 Grigore, George, 7, 26, 29, 30, 35, 36, 38
 Grosjean, François, 328
 Grotzfeld, Heinz, 312
 Gulle, Ozan, 48, 55, 56
 Gurlekian, Jorge, 390, 391
 Gussenhoven, Carlos, 395, 397

 Haase, Martin, 325
 Hachimi, Atiqa, 309
 Haedo, Diego de, 289, 291, 297–299
 Haeri, Niloofar, 306, 309
 Haig, Geoffrey, 34
 Hall, Robert A. Jr, 287
 Hamid Ahmed, Mohamed-Tahir, 208, 209, 211
 Hamza, Belgacem, 240
 Hansen, Anita Berit, 394
 Hansen, Maj-Britt Mosegaard, 323, 365, 366
 Harrell, Richard S., 96
 Haser, Verena, 153
 Haspelmath, Martin, 118, 324, 378
 Hassan, Jidda, 137, 138, 143, 152
 Hassan, Zeki Majeed, 314
 Heath, Jeffrey, 66, 67, 84, 86, 88, 89, 91, 100, 334, 395
 Heath, Jeffrey., 74, 85, 92
 Heim, Irene, 369
 Heine, Bernd, 138, 147, 323–325, 329, 331, 332, 349, 351, 376
 Hellmuth, Sam, 389, 393, 394, 396–398
 Hellmuth, Samantha, 104
 Herin, Bruno, 104, 266, 284, 285, 309, 313, 394
 Heselwood, Barry, 251, 314
 Holes, Clive, 6, 309, 398
 Hopkins, J. F. P., 299
 Hopper, Paul, 322
 Horesh, Uri, 305, 309, 313
 Hoyland, Robert G., 246
 Hudson, Richard A., 211
 Hutchison, John, 142, 148

 Ibrahim, Muhammad H., 306
 Iemmolo, Giorgio, 114, 115
 Il-Hazmy, Alayan Mohammed, 331
 Ingham, Bruce, 4, 7, 15, 17, 18, 35
 Iraqui Sinaceur, Zakia, 90
 Isaksson, Bo, 26, 42

 Jastrow, Otto, 23–31, 35–38, 40, 42, 43, 333
 Jenkins, Edward V., 165
 Jespersen, Otto, 362
 Johnson, Mark, 153
 Johnstone, T. M., 373
 Jun, Sun-Ah, 388, 395, 396

 Kahane, Henry, 287, 288
 Kahane, Renée, 287, 288
 Kahrel, Peter, 363
 Kariolemou, Marilena, 48
 Kaufman, Terrence, 349, 389
 Keane, G. J., 165
 Keesing, Roger, 356
 Kerswill, Paul, 311

Name Name index

- Khan, Geoffrey, 114, 116, 124
 Khat tab, Ghada, 314
 Kherbache, Fatma, 239
 Kireva, Elena, 104
 Klimiuk, Maciej, 101
 Knecht, 295
 Kogan, Leonid, 246
 König, Ekkehard, 351
 Kossmann, Maarten, 68–71, 73, 77, 233,
 235–240, 351, 375
 Kramer, Raija, 153
 Krapova, Iliyana, 115
 Krier, Fernande, 125
 Kulk, Friso, 398
 Kuteva, Tania, 138, 147, 325, 329, 331,
 332, 349, 376

 Ladd, D. Robert, 388, 389
 Ladefoged, Peter, 250, 251
 Ladusaw, William A., 364
 Lafkioui, Mena, 71, 375, 378
 Lahdo, Ablahad, 26, 31–35, 38, 42, 117
 Laka, Itziar, 366
 Lakoff, George, 153
 Lash, Elliott, 362, 371–373
 Law, Daniel, 325, 326
 Le Page, R. B., 311, 328
 Leddy-Cecere, Thomas, 147², 305, 328,
 329, 336, 337, 349, 376
 Lefebvre, Claire, 354
 Lehmann, Christian, 323, 324, 327
 Leitner, Bettina, 378
 Lentin, Jérôme, 312, 331
 Levin, Aryeh, 312
 Lewin, Bernhard., 331
 Lichtenberk, Frantisek, 363
 Lipiński, Edward, 124
 Lonnet, Antoine, 246, 250–253, 257,
 258, 260

 Loporcaro, Michele, 102
 Lucas, Christopher, 6, 12, 15, 18, 28,
 37, 55, 58, 65, 71, 73, 89–91,
 99, 153, 160, 191, 192³, 194,
 196, 199, 238, 249, 305, 345,
 352, 362, 366, 367, 371–373,
 375, 379, 387
 Luffin, Xavier, 161, 162, 167, 169, 170,
 173, 174, 181
 Lydon, Ghislaine, 335

 Maas, Utz, 395
 Macalister, Robert Alexander Stew-
 art, 266
 Macúch, R., 124
 Maddieson, Ian, 250, 251
 Mahdi, Qasim R., 330
 Maiden, Martin, 110, 111
 Majidi, Mohammad-Reza, 7, 9, 12, 14
 Malaika, Nisar, 17
 Mammeri, Mouloud, 231
 Manfredi, Stefano, 85, 112, 115, 116, 136,
 159, 160, 163, 164, 166–169,
 174, 192, 208¹, 211, 215, 275,
 305, 327, 348, 351–353, 376
 Marçais, Philippe., 70, 71
 Matras, Yaron, 4, 5, 7, 8, 10, 16, 17, 19,
 76, 198, 266, 269, 270, 284,
 285, 337, 376, 377, 391
 Mechehed, Djamel-Eddine, 233
 Meillet, Antoine, 361
 Meldon, J. A., 165
 Mennen, Ineke, 389
 Meouak, Mohamed, 240
 Mettouchi, Amina, 351
 Meyerhoff, Miriam, 356
 Mifsud, Manwel, 99, 106–110, 118, 119,
 125

Name Name index

- Miller, Catherine, 160, 166, 168, 172, 181, 199, 321
- Minervini, Laura, 287, 288
- Mion, Giuliano, 66
- Moi, Daniel R., 171
- Mori, Laura, 288
- Morin, Didier, 208, 211
- Morris, Miranda J., 248, 253, 260
- Morton, Rachel, 389
- Mouili, Fatiha, 230
- Mourigh, Khalid, 236, 237, 240, 241
- Mughazy, Mustafa, 370
- Muraz, Gaston, 161
- Múrcia Sánchez, Carles, 229
- Mufāwiya, 47
- Myers-Scotton, Carol, 76
- Naciri-Azzouz, Amina, 69
- Næss, Unn Gyda, 177, 178
- Naïm, Samia, 312
- Naït-Zerrad, Kamal, 100, 123
- Nakano, Aki'o, 146
- Nakao, Shuichiro, 167, 168, 170–174, 181, 348, 349, 397
- Nance, Claire, 390
- Naumann, Christfried, 235
- Naumkin, Vitaly, 251
- Nebel, Arthur, 170
- Newton, Brian, 54–56, 58, 59
- Nolan, Joanna, 67, 84, 299
- Norde, Muriel, 324
- Nortier, Jacomine, 92
- Nycz, Jennifer, 307
- O'Rourke, Erin, 104, 390
- Oberling, P., 3, 4
- Ondráčková, Zuzana, 123
- Onour, ʿAbdallah, 208
- Öpengin, Ergin, 27, 34
- Owen, H. R., 165
- Owen, Roger, 353
- Owens, Jonathan, 28, 136–138, 140, 141, 143, 151, 152, 161, 163, 165–167, 172, 192, 192³, 200, 346–348
- Palva, Heikki, 35
- Pananti, Filippo, 291, 296, 299
- Panza, G., 171
- Papaconstantinou, Arietta, 396, 397
- Paradisi, Umberto, 238, 375
- Pat-El, Na'ama, 10
- Pat-El, Na'ama, 367
- Pellegrini, Giovanni Battista, 293
- Pennell, Richard, 295, 296
- Pepper, Stephen, 344, 345
- Pereira, Christophe, 66
- Persson, Maria, 254
- Peust, Carsten, 397
- Piccitto, Giorgio, 113
- Pichler, Werner, 231
- Pihan, Antoine Paulin, 297
- Plantet, Eugène, 290
- Poiret, Jean Louis Marie, 296
- Poplack, Shana, 198
- Porkhomovsky, Victor, 251
- Power, Timothy, 335
- Pozzati, Aurelio, 171
- Prieto, Pilar, 389, 397
- PRIO, 50
- Procházka, Stephan, 4, 6, 7, 10, 23–25, 27, 30, 34, 42, 48, 74, 331, 395
- Qafisheh, Hamdi, 331
- Queen, Robin M., 389
- Rabeh, 138, 141, 161

Name Name index

- Rabhi, Allaoua, 375, 378
 Rahmani, Hamed, 395
 Ramat, Paolo, 323, 362
 Rastegar-El Zarka, Dina, 396
 Ratcliffe, Robert, 13, 198, 354
 Ratcliffe, Robert R., 193
 Rehbinder, Johan von, 296
 Reinhardt, Carl, 333
 Retsö, Jan, 138, 246
 Ridouane, Rachid, 250
 Rifaat, Khalid, 396
 Rilly, Claude, 207
 Ritt-Benmimoun, Veronika, 111, 150, 347, 348
 Robin, Christian-Julien, 246
 Robustelli, Cecilia, 110, 111
 Roger II of Sicily, 97
 Romaine, Suzanne, 191–193, 200
 Roper, E. M., 211
 Rosenhouse, Judith, 312, 330
 Ross, Malcolm, 327, 343, 344, 349, 355
 Rossetti, Roberto, 292
 Roth, Arlette, 51, 54, 57–59
 Rouchdy, Aleya, 188, 190, 200
 Rubin, Aaron D., 246, 324
 Russi, Cinzia, 116
 Saada, Lucienne, 332
 Saade, Benjamin, 107
 Sadler, Louisa, 117, 379
 Saltarelli, Mario, 190, 191, 197, 200
 Sammut, Carmen, 121
 Sánchez, Pablo, 332
 Sankoff, Gillian, 187
 Sasse, Hans-Jürgen, 344
 Savary de Brèves, François, 290
 Sayahi, Lotfi, 71, 74, 77, 85, 92, 394
 Schilling-Estes, Natalie, 334
 Schroepfer, Jason, 332
 Schuchardt, Hugo, 293, 295, 296, 299
 Sciriha, Lydia, 99
 Seeger, Ulrich, 192, 367
 Seifart, Frank, 105
 Selbach, Rachel, 291, 296
 Serreli, Valentina, 230
 Shabibi, Maryam, 4, 5, 7, 8, 10, 16, 17, 19
 Shahin, Kimary, 309, 311–313
 Silverstein, Michael, 328
 Sima, Alexander, 259
 Simeone-Senelle, Marie-Claude, 213, 247, 253, 257, 260, 368
 Simonet, Miquel, 390
 Smart, Jack R., 175, 179
 Sokoloff, Michael, 124
 Sorace, Antonella, 391
 Souag, Lameen, 67–69, 71, 75, 77, 85, 116, 123, 233, 235–241, 354, 367, 375–378
 Spagnol, Michael, 105, 116, 118–120, 125, 379
 Stassen, Leon, 270
 Stephens, Laurence D., 397
 Stevenson, R. C., 149
 Stewart, Devin, 337
 Stolz, Thomas, 125
 Sweetser, Eve, 153
 Tabouret-Keller, Andrée, 311, 328
 Tadmor, Uri, 118, 239
 Taine-Cheikh, Catherine, 69, 70, 87, 96, 332, 345, 346, 350, 351
 Talay, Shabo, 26, 28, 30, 35, 38–40, 43
 Talmoudi, Fathi., 70, 74, 330
 Taylan, Eser, 40, 41
 Thomason, Sarah G., 26, 343, 349, 389
 Tigziri, Nora, 230
 Tinniswood, Adrian, 290

Name Name index

- Tosco, Mauro, 159, 161, 163, 164, 174
 Traugott, Elizabeth, 322
 Trudgill, Peter, 160, 306, 307, 310–312
 Tsiapera, Maria, 51, 54, 56
 Tully, Miss, 295, 297
 Türkmen, Erkan, 75
- Ullendorf, Edward, 213
 Ullmann, Manfred, 112
 Ursini, Flavia, 294
- Van Coetsem, Frans, 26, 68, 74, 76, 99,
 137, 160, 168, 191, 199, 249,
 344, 372
 van der Auwera, Johan, 361, 362, 371
 van der Wal Anonby, Christina, 373,
 374
 van Putten, Marijn, 235, 237, 240
 Vanhove, Martine, 97, 100, 112, 117,
 124, 209–212, 217, 219, 220,
 223, 330, 352
 Vassalli, Michelantonio, 96
 Vassallo, Mario, 99
 Vella, Alexandra, 104
 Versteegh, Kees, 13, 24, 25, 38, 108,
 190, 192, 246
 Vicente, Ángeles, 74, 84, 190
 Visconti, Jacqueline, 323
 Vocke, Sybille, 30
 Vollers, Karl, 324, 325
 Vossen, Frens, 362, 371
 Vycichl, Werner, 235
- Waldner, Wolfram, 30
 Walker, Traci, 391
 Walter, Mary-Ann, 24, 28, 37, 56, 101,
 102, 117, 192³
 Walters, Keith, 394
 Warrington, 299
- Watson, Janet C. E., 246, 251, 254, 255,
 257, 259, 260, 332, 333, 368,
 395, 397
 Wedekind, Klaus, 208, 210, 223
 Weinreich, Uriel, 389
 Weiss, Gillian, 290
 Wellens, Inneke, 165, 167–169, 171–173
 Wells, John C., 388
 Wellsted, James R., 245
 Williams, Ann, 311
 Willis, David, 362, 364, 380
 Wilmsen, David, 345, 367, 381
 Winford, Donald, 7, 74, 160, 180, 249,
 347, 352
 Wittrich, Michaela, 26, 30
 Wohlgemuth, Jan, 222
 Woidich, Manfred, 16, 101, 146, 312,
 321, 345
 Wolfram, Walt, 334
 Woodhead, D. R., 13
- Yoda, Sumikazu, 67, 75
- Zaborski, Andrzej, 219
 Zagona, Karen, 116
 Záhořík, Jan, 209
 Zammit, Martin R., 104, 246, 394
 Ziamari, Karima, 74, 77
 Zibelius-Chen, Karola, 207
 Zuckermann, Ghil’ad, 356

Language index

CyA, 55–58

Cypriot Greek, 58

English

Maltese, 99

Kurdish, 40, 41

Lebanese Arabic, 58

Sason Arabic, 40¹⁴

Sicilian, 113

Siwi, 238

Tillo Arabic, 32, 33

Did you like this book?

This book was brought to you for free

Please help us in providing free access to linguistic research worldwide. Visit <http://www.langsci-press.org/donate> to provide financial support or register as a community proofreader or typesetter at <http://www.langsci-press.org/register>.



Arabic and contact-induced change

change blurb in localmetadata.tex

