Arabic and contact-induced change

Edited by

Christopher Lucas

Stefano Manfredi

Contact and Multilingualism

language science press

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

ISSN:

Arabic and contact-induced change

Edited by

Christopher Lucas

Stefano Manfredi

language
science
press

Lucas, Christopher & Stefano Manfredi (eds.). 2020. *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

ISBN: no digital ISBN
no print ISBNs!
ISSN:
no DOI
ID not assigned!

Cover and concept of design: Ulrike Harbort
Fonts: Libertinus, Arimo, DejaVu Sans Mono, SIL Scheherazade
Typesetting software: XeLaTeX

Language Science Press
Xhain
Grünberger Str. 16
10243 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

Contents

Preface

Acknowledgments

Part I

Contact-induced change in varieties of Arabic

Chapter 1

Pre-Islamic Arabic

Ahmad Al-Jallad
The Ohio State University

This chapter provides an overview of Arabic in contact in the pre-Islamic period, from the early first millennium BCE to the rise of Islam. Contact languages include Akkadian, Aramaic, Ancient South Arabian, Canaanite, Dadanitic, and Greek. The chapter concludes with two case studies on contact-induced development: the emergence of the definite article and the realization of the feminine ending.

1 Language contact in the pre-Islamic period

> [I]the Djāhiliyya, 'the Age of Ignorance' [...], the Arabs lived to a great extent in almost complete isolation from the outer world... [t]his accounts for the prima facie astonishing fact that Arabic, though appearing on the stage of history hundreds of years after the Canaanites and Aramaeans, nevertheless in many respects has a more archaic character than these old Semitic languages. The Arabs, being almost completely isolated from outer influences and living under the same primitive conditions of their ancestors preserved the archaic structure of their language. (Blau 1981: 18).

This is the image of Arabic's pre-Islamic past that emerges from Classical Arabic sources. For writers such as Ibn Khaldūn, contact-induced change in Arabic was a by-product of the Arab conquests, and served to explain the differences between the colloquial(s) of his time and the literary language. More than a century and a half of epigraphic and archaeological research in Arabia and adjacent areas has rendered this view of Arabic's past untenable. Arabic first appears in the epigraphic record in the early first millennium BCE, and for most of its pre-Islamic history, the language interacted in diverse ways with a number of related

Ahmad Al-Jallad. 2020. Pre-Islamic Arabic. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 3–21. Berlin: Language Science Press. DOI:??

Semitic languages and Greek. This chapter will outline the various foci of contact between Arabic and other languages in the pre-Islamic period based on documentary evidence. Following this, I offer two short case studies showing how contact-induced change in the pre-Islamic period may explain some of the key features of Arabic today.

1.1 Old Arabic

Old Arabic is an umbrella term for the diverse forms of the language attested in documentary and literary sources from the pre-Islamic period, including inscriptions, papyri, and transcriptions in Greek, Latin, and cuneiform texts. The present usage does not refer to Classical Arabic or the linguistic material attributed to the pre-Islamic period collected in the eighth and ninth centuries CE, such as poetry and proverbs, as we cannot be sure about their authenticity, especially with regard to their linguistic features. Al-Jallad (2017) defines the corpus of Old Arabic as follows: Safaitic, an Ancient North Arabian script concentrated in the Syro-Jordanian Ḥarrah (end of the 1st millennium BCE to 4th c. CE), Hismaic, an Ancient North Arabian script spanning from central Jordan to northwest Arabia (chronology unclear, but overlapping with Nabataean), the substratum of Nabataean Aramaic, along with a few Arabic-language texts carved in this script (2nd c. BCE to 4th c. CE), the Nabataeo-Arabic inscriptions (3rd c. CE to 5th c. CE), pre-Islamic Arabic script inscriptions (5th c. CE to early 7th c. CE) and isolated inscriptions in the Greek, Dadanitic (the oasis of Dadān, modern day al-ʕUlā, northwest Ḥiǧāz), and Ancient South Arabian alphabets (varied chronology).

In geographic terms, Old Arabic is attested mainly in the southern Levant, the Sinai, and northwestern Arabia, as far south as Ḥegrā (Madāʔin Ṣāleḥ). Within this area a variety of non-Arabic languages were spoken and written, with which Old Arabic interacted. The main contact language was Imperial Aramaic, which served as a literary language across North Arabia in the latter half of the first millennium BCE until, perhaps, the rise of Islam. Since contact must be viewed through the lens of writing, it is in most cases difficult to determine how extensive multilingualism was outside of literate circles.

2 Contact languages

2.1 Arabic and Akkadian

The first attestations of Arabic are preserved in cuneiform documents. While no Arabic texts written in cuneiform have yet been discovered, isolated lexical items survive in this medium. Livingstone (1997) identified an example of the Old Arabic word for 'camel' with the definite article in the inscriptions of Tiglath-pileser III (744–727 BCE): *a-na-qa-a-te* = *(h/ʔ)an-nāq-āte* 'the she-camels'. Aside from this, almost all other Arabic material consists of personal and divine names. There are reports of "Arabs" in Mesopotamia – inhabiting walled towns in western Babylonia – as early as the eighth century BCE (Ephʿal 1974: 112). While we cannot be sure that the people whom the Babylonians called Arabs were in fact Arabic speakers, a few texts in dispersed Ancient North Arabian scripts hail from this region. So far, all seem to contain only personal names with Arabic or Arabian etymologies.[1] These facts can only suggest the possibility of contact between speakers of Arabic and Akkadian in the early first millennium BCE.

2.2 Arabic and Canaanite

Contact between Arabic speakers and speakers of Canaanite languages is documented in the Hebrew Bible (Ephʿal 1982: ch.2; Retsö 2003: ch.8), and there is one inscription directly attesting to contact between both groups. An Ancient North Arabian inscription from Bāyir, Jordan contains a prayer in Old Arabic to three gods of the Iron Age Canaanite kingdoms of Moab, Ammon, and Edom (Hayajneh et al. 2015). The text is accompanied by a Canaanite inscription, which remains undeciphered. The reading of the Arabic according to the edition is as follows:

(1) Bāyir inscription (Hayajneh et al. 2015)
h mlkm w-kms w-qws b-km ʕwðn h-ʔsḥy
VOC PN CONJ-PN CONJ-PN PREP-2PL.M protect.PRF.1PL DEM-well.PL
m-mdwst
PREP-ruin
'O Malkom, Kemosh, and Qaws, we place under your protection these wells against ruin.'

[1] "Dispersed Ancient North Arabian" is a temporary term given to the Ancient North Arabian inscriptions on seals, pottery, bricks, etc. which have been found in various parts of Mesopotamia and elsewhere (Macdonald 2000: 33).

2.3 Arabic and Aramaic

Evidence for contact between Arabic and Aramaic spans from the middle of the first millennium BCE to the late sixth century CE, and is concentrated in the southern Levant and northwest Arabia.[2] Perhaps one of the earliest examples of Arabic speakers using Aramaic as a written language comes from the fifth-century-BCE Nile Delta. A king of Qedar, Qayno son of Gośam,[3] commissioned an Aramaic votive inscription dedicated to *hn-ʔlt* 'the goddess' (Rabinowitz 1956). Arabic names can be found in transcription across the Levant in Aramaic inscriptions (Israel 1995), and in most cases names with an Arabic etymology terminating in the characteristic final *-w*, reflecting an original nominative case (Al-Jallad forthcoming).

Arabic and Aramaic language contact reaches a climax in the written record at the end of the first millennium BCE with the arrival of inscriptions in the Nabataean script. The Nabataeans established a kingdom in the region of Edom in the fourth century BCE, which at its greatest extent spanned from the Ḥawrān to the northern Ḥiğāz. While they, like their contemporaries across the Near East, wrote in a form of Imperial Aramaic, the spoken language of the royal house and large segments of the population was Arabic. Unlike other examples of Aramaic written by Arabic speakers so far, Nabataean incorporated Arabic elements into its writing school, such as the optative use of the perfect, the negator *ɣayr*, and a significant number of lexical items relating to daily life (Gzella 2015: 242–243).

Perhaps one of the most interesting examples of contact between the two languages is found in Nabataean legal papyri from the Judaean desert (1st–2nd c. CE). These Aramaic-language legal documents contain a number of glosses in Arabic, for example: *ʕqd* /ʕaqd/ 'contract'; *mʃnm* /maɣnam/ 'profit'; *prʃ* /faraʕ/ 'to branch out'; *ṣnʕh* /ṣanʕah/ 'handiwork', etc. (Yardeni 2014). Macdonald (2010: 20) has suggested, based on this evidence, that Nabataean legal proceedings would have taken place in Arabic, while all written records were made in Aramaic.

In addition to the use of Arabic within Aramaic, a unique votive inscription from ʕEn ʕAvdat (Negev, Israel) contains three verses of an Arabic hymn to the deified Nabataean king ʕObodat embedded within an Aramaic text. While undated (but likely earlier than 150 CE), the text is certainly the earliest example of continuous Arabic language written in the Nabataean script, as before this almost all examples are isolated words and personal names.

[2] See Stein (2018) on the role of Aramaic in the Arabian Peninsula in the pre-Islamic period.

[3] The symbol *ś* denotes the Old Arabic reflex of Classical Arabic ⟨ش⟩, which is usually transcribed *š*. /ś/ was likely realized as a voiceless lateral fricative [ɬ].

(2) ʕEn ʕAvdat inscription[4]

a. Aramaic
dkyr b-ṭb q[r]ʔ qdm ʕbdt ʔlhʔ
remember.PTCP.PASS PREP-good read.PTCP.ACT PREP PN god.DEF
w-dkyr mn ktb grmʔlhy br
CONJ-remember.PTCP.PASS REL write.PRF.3SG.M PN son.CS
tymʔlhy šlm lqbl ʕbdt ʔlhʔ
PN be_secure.PRF.3SG.M PREP PN god.DEF
'May he who reads this aloud be remembered for good before ʕObodat the god, and may he who wrote be remembered. May Garmallāhi son of Taymallāhi be secure in the presence of ʕObodat (the god).'

b. Arabic
p-ypʕl lʔ pdʔ w-lʔ ʔtrʔ p-kn
CONJ-act.IMPF.3SG.M NEG ransom.ACC CONJ-NEG scar.ACC CONJ-be.INF
hnʔ ybʕ-nʔ ʔl-mwtw lʔ ʔbʕ-h
here seek.IMPF.3SG.M-1PL DEF-death.NOM NEG make.obtain.INF-3SG.M
p-kn hnʔ ʔrd grḥw lʔ
CONJ-be.INF here want.PRF.3SG.M wound.NOM NEG
yrd-nʔ
want.IMPF.3SG.M-1PL
'May he act that there be neither ransom nor scar; so be it that death would seek us, may he not aid its seeking, and so be it that a wound would desire (a victim), let it not desire us!'

c. Aramaic
grmʔlhy ktb yd-h
PN writing.CS hand.3SG.M
'Garmallāhi, the writing of his hand.'

The presence of Aramaic is much more lightly felt in the desert hinterland to the east and north. A small handful of Safaitic–Aramaic bilingual inscriptions are known (Hayajneh 2009: 214–215). In one Safaitic text, produced by a Nabataean, the author gives his name and affiliation to social groups in a type of Aramaic, but then writes the remainder of the inscription in Old Arabic, suggesting that this individual may have been bilingual.

[4]This is my translation; the *editio princeps* is Negev, Naveh & Shaked (1986); it is discussed most recently in Fiema et al. (2015: 399–402) and Kropp (2017).

(3) Nabataean Safaitic (Al-Jallad 2015: 19; C 2820)
l ʔʔsd bn rbʔl bn ʔʔsd bn rbʔl nbṭwy slmy w
PREP PN son.CS PN son.CS PN son.CS PN Nabataean Salamite CONJ
brḥ ḫlqt śty h-dr w
depart.PRF.3SG.M period.CS winter DEF-region CONJ
tḏr h-smy
keep_watch.PRF.3SG.M DEF-sky
'By ʔAʔsad son of Rabbʔel son of ʔAʔsad son of Rabbʔel, the Nabataean Salamite, and he set off from this place for the period of winter and kept watch for the rains.'

A handful of Aramaic loans are found in the Safaitic inscriptions: *sfr* 'writing'; *ʔsyt* 'hide, trap', *lṣṭ* 'thief', ultimately from Greek *lēistḗs*. Other words, such as *mdbr* /madbar/ 'the Hamad, wilderness' and *nḫl* /naḫl/ 'valley', are absent in Classical Arabic yet appear in the Northwest Semitic languages. These do not appear to be loans, however, as their meanings and phonologies are local and Arabic, respectively. They should instead be regarded as genuine cognates that did not make it into the Islamic-period lexica.

2.4 *Provincia Arabia* and the Nabataeo-Arabic script

In 106 CE, under circumstances that remain poorly understood, the Romans annexed the Nabataean Kingdom and established their Province of Arabia. While Nabataean political independence ended, their script, writing tradition and language continued to thrive and evolve. This is exemplified by the famous tomb inscription of Raqōś bint ʕAbd-Manōto from Madāʔin Ṣāliḥ. Dated to 267 CE, the text is a legal inscription associated with the grave of a woman who died in al-Ḥegr. Unlike other grave inscriptions at this site, the Raqōś inscription is composed almost entirely in Arabic, with the Aramaic components restricted to the introductory demonstrative *dnh* 'this', the words for 'son' and 'daughter', the dating formula, and the name of the deity. The Aramaic components are bolded below:

(4) Madāʔin Ṣāliḥ inscription (JSNab 17)[5]
dnh qbrw ṣnʕ-h kʕbw **br** ḥrtt l-rqwš **brt**
DEM grave build.PRF.3SG.M-3SG.M PN son PN PREP-PN daughter

[5]For the latest discussion of this text, see Macdonald's contribution to Fiema et al. (2015: 402–405).

ʕbdmnwtw ʔm-h w-h hlkt fy ʔl-ḥgrwy **šnt**
PN mother-3SG.M CONJ-3SG.F die.PRF.3SG.F PREP DEF-PN year
mʔh **w-štyn** **w-tryn** **b-yrḥ** **tmwz**
hundred CONJ-sixty CONJ-two PREP-month Tammūz
w-lʕn **mryʕlmʔ** mn yšnʔ ʔl-qbrw d[ʔ]
CONJ-curse.PRF.3SG.M PN REL desecrate.IMPF.3SG.M DEF-grave DEM
w-mn yftḥ-h ḥšy w wld-h
CONJ-REL open.IMPF.3SG.M-3SG.M except w children-3SG.M
w-lʕn mn yqbr w-yʕly
CONJ-curse.PRF.3SG.M REL bury.IMPF.3SG.M CONJ-remove.IMPF.3SG.M
mn-h
PREP-3SG.M

‘**This** is a grave that KaʕboSON of Ḥāreθat constructed for Raqōś **daughter** of ʕAbd-Manōto, his mother, and she perished in al-Ḥegro **year one hundred and sixty two in the month of Tammūz**. May **the Lord of the World** curse anyone who desecrates this grave and anyone who would open it, with the exception of his children, and may he curse anyone who would bury or remove from it (a body).’

During the same period, the classical Nabataean script continues to evolve towards what we consider the Arabic script (Nehmé 2010). Its letter forms take on a more cursive character, and the connecting element of each letter goes across the bottom of the text. Nehmé considers the letter forms typical of the Arabic script to have evolved from Nabataean between the third and fifth centuries CE. In inscriptions from this period, the Arabic component begins to increase at the expense of Aramaic (Nehmé 2017). This trend may suggest that knowledge of Aramaic was waning in these centuries, or that the writing tradition itself was transforming – Aramaic was slowly being replaced by Arabic. If we think in terms of writing schools, there may not have been much Arabic–Aramaic bilingualism in Arabia outside of the scribal class – indeed, scholars have continued to debate whether Nabataean Aramaic was ever a colloquial, and there are good arguments to doubt that it was (Gzella 2015: 240). The remnants of Aramaic in the latest phases of the Nabataeo-Arabic inscriptions, however, most certainly functioned as a code, grams for Arabic words, a situation comparable to the Aramaeograms of Pahlavi (cf. Nyberg 1974).

2.5 The Arabic inscriptions of the sixth century CE

In Arabic inscriptions of the sixth century, written Arabic and Aramaic continue the stable situation of contact witnessed in the Nabataeo-Arabic period. Aramaic fossils are employed in dating formulae and the word for 'son', and possibly the first person pronoun. But otherwise, the language of these texts is entirely Arabic. Perhaps the most famous among these is the inscription of Jebel Usays, given in (5), in which the Aramaic components are bolded.

(5) Jebel Usays inscription[6]
ʔnh[7] rqym **br** mʕrf ʔl-ʔwsy ʔrsl-ny ʔlḥrt ʔl-mlk ʕly
1SG PN son PN DEF-Awsite send.PRF.3SG.M-1SG PN DEF-king PREP
ʔsys mslḥh snt 423
Usays outpost year 423
'**I**, Ruqaym **son** of Muʕarrif the Awsite al-Ḥāriθ the king sent me to Usays as an outpost, year 423.' [= 528/9 CE]

2.6 Arabic, Greek and Aramaic in sixth-century Petra

In 1993, a corpus of carbonized Greek papyri – some 140 rolls – was discovered at the Byzantine church of Petra.[8] These documents attest to a trilingual situation at the city: Greek served as the official administrative language, while Arabic and Aramaic appear to have been spoken languages. The microtoponyms (names of small plots of lands and vineyards) are in both Arabic and Aramaic, and often times the same word is expressed in both languages, as in Table (1).

This naturally suggests that, alongside literacy in Greek, there was spoken bilingualism in Arabic and Aramaic, perhaps a stable situation extending back to Nabataean times.

2.7 Arabic and Ancient South Arabian

Classical Arabic sources note a situation of close contact between Arabic and "Ḥimyaritic", a term used for a language they associated with the pre-Islamic

[6]For the latest discussion of this text, see Macdonald's contribution to Fiema et al. (2015: 405).

[7]While it has been suggested that the spelling *ʔnh* reflects a pausal form (Larcher 2010), it seems more likely in light of the Thaʕlabah Nabataeo-Arabic inscription (Avner et al. 2013), which spells 'I' as *ʔnh*, that this form reflects the Aramaic spelling of the pronoun rather than an Arabic variant.

[8]These papyri are edited in a five-volume series: the *Petra Papyri I–V* (2002–2018), various editors, Amman: American Center of Oriental Research. See Arjava et al. (2018) for the last volume.

Table 1: Arabic–Aramaic equivalents in the Petra Papyri (Al-Jallad 2018a: 41)

Translation	Arabic	Aramaic
'land markers'	Αραμ /ārām/	Εραμαειια /eramayyā/
'farm'	αλ-Ναcβα /al-naṣbah/	Ναcβαθα /naṣbatā/
'canal'	αλ-Κεcεβ /al-qeṣeb/	Κιcβα/Κειcβα /qiṣbā/

kingdom of Ḥimyar in what is today Yemen. The pre-Islamic inscriptions from the northern Yemeni Jawf, the so-called Haram region, attest to a similar situation. These texts are composed in Sabaic, but contain a significant admixture of non-Sabaic linguistic material. Some scholars (e.g. Robin 1991) have considered Arabic to be the contributing source, but in most cases the non-Sabaic linguistic features are not specific to Arabic, such as the use of the causative verb ʔaCCaC, which is attested in Aramaic and Ge'ez for example, rather than haCCaC as in Sabaic. As Macdonald (2000: 55) rightly puts it, these inscriptions are basically Sabaic, with a small admixture from North Arabian languages, but not necessarily Arabic. Four texts from this region, however, exhibit the Arabic isogloss of *lam* for past-tense negation, suggesting that some form of Arabic may have contributed to their mixed character.[9]

Mixed North/South Arabian texts can be found further to the north, in Nağrān and Qaryat al-Fāw. The most famous is perhaps the grave inscription of *Rbbl bn Hfʕm*. This unique text attests features that can be attributed to both non-Sabaic and Sabaic sources. On the non-Sabaic side, it uses the definite article *ʔl*, the causative morpheme *ʔ-* rather than *h-*, and occasionally the 3rd person pronoun *h* rather than *hw*. At the same time, the text employs mimation, clitic pronouns with long vowels, e.g. *-hw*, and prepositions not known in Arabic (Al-Jallad 2018b: 30). At Nağrān, one occasionally encounters Arabic lexical items, such as *ldy* 'at' and *ʕnd* 'with' in otherwise perfectly good South Arabian texts. So then, how are we to interpret the mixed character of these texts? For Qaryat al-Fāw, Durand (2017: 95, fn.32) has suggested, based on the significant amount of Petraean pottery, that a sizable Nabataean colony existed at the oasis. It could be the case that Nabataean colonists introduced Arabic to the oasis, where it naturally gained prestige as a trade language given its links with the north. The mixed nature of some of the inscriptions of this site could therefore be interpreted in

[9]For a list of the Haram inscriptions, see Macdonald (2000: 61), who labels these texts Sabaeo-North-Arabian.

two ways. If they reflect a spoken variety, then perhaps they are the result of convergence between the Arabic introduced by the Nabataeans and Sabaic, similar to the modern dialects of Yemeni Arabic today, which are essentially Arabic with a significant South Arabian admixture.[10] If we are dealing with an artificial scribal register, then the language may be the result of a scribe attempting to produce a text in Arabic, for an Arabic-speaking customer, but inadvertently introducing Sabaicisms from the language he is more used to writing. A similar phenomenon might be at play in the Aramaic–Hasaitic tomb inscription from Mleiha.[11] There, the scribe – seemingly unintentionally – uses the Aramaic word for son, *br*, in the Hasaitic portion of the text, suggesting perhaps that he was bilingual and more used to writing in Aramaic (Overlaet et al. 2016).

2.8 Arabic in the Ḥiǧāz

Before the arrival of the Nabataeans, the written language of the oasis of al-ʕUlā and associated environs in the northern Ḥiǧāz was Dadanitic, a non-Arabic Central Semitic language. A few texts, however, display features that are unambiguously Arabic. The best known of these is JSLih 384. This short text is written in the Dadanitic script but seems to be, in other respects, produced in a dialect of Old Arabic, notably making use of the relative pronoun *ʔlt* /ʔallatī/. Two other Dadanitic texts make use of the Arabic construction *ʔn yfʕl*, that is, the use of the subordinator *ʔan* with a modal verb. In addition to this, one occasionally finds the *ʔ(l)* definite article employed in these inscriptions. The interpretation of this contact situation, like that in South Arabia, is unclear. Do these few texts represent the writings of travelers or immigrants from the north, whose spoken language influenced the dictation of text to the scribe? Or do they reflect unique points on a dialect continuum? The complex linguistic situation at ancient Dadan is the subject of a fascinating study by Kootstra (2019).

2.9 Arabic and the languages of the Thamudic inscriptions

Even more difficult to distill is the possible contact situation between Arabic and the more shadowy pre-Arabic Semitic languages of north and central Arabia. We are afforded a small glimpse of these languages by the laconic Thamudic inscriptions, mainly those classified in the C, D, and F scripts.[12] While it is difficult to

[10] On these varieties, see Watson (2018).

[11] Hasaitic is the name given to the pre-Islamic script and language of East Arabia.

[12] Thamudic B, C, and D are discussed in Macdonald (2000) and Al-Jallad (2017; 2018b); Thamudic F is outlined in Prioletta & Robin (2018).

say much about the languages these scripts express, they are clearly distinct from Arabic (Al-Jallad 2017: 321–322). The only evidence for contact between Arabic and any of these languages is found in the tomb inscription of Raqōś at Madāʔin Ṣāliḥ, illustrated in (4). This text, as discussed in §2.4, is written mainly in Arabic, with a few fossilized Aramaic components. Alongside the main inscription, there is a short text inscribed in the Thamudic D script stating: *ʔn rqś bnt ʕbdmnt* 'This is Raqōś, daughter of ʕAbdo-Manōto'. The use of the introductory element *ʔn* 'this' or perhaps 'for', rather than the Arabic demonstrative *dʔ* /ðā/ or perhaps its feminine equivalent *dy* /ðī/, employed in the Nabataean text, indicates that we are dealing with a third language.[13] Did Raqōś originally hail from a nomadic community who spoke a non-Arabic Semitic language expressed in the Thamudic D script? And did she later come to live in Arabic-speaking Ḥegrā? Was the use of this script on her grave a tribute to her heritage? These questions are impossible to answer with the data available to us now, but they widen the scope of investigation when examining Arabic's history. The available fragments of evidence support the suggestion put forth recently by Souag (2018): we must consider the possibility of unknown Semitic substrate(s) in the development of early Arabic.

2.10 Arabic and Greek

The nexus of Arabic–Greek contact, based on the inscriptions known so far, is the Syro-Jordanian Harrah, the basalt desert that spans from the Hawrān to northern Arabia. Greek inscriptions are occasionally found throughout this region, interacting with the local Arabic dialects in diverse ways. The commonest type of bilingual text consists of simple signatures in Safaitic and Greek. These texts, illustrated in (6), only prove that the author knew how to write his name in Greek, and do not constitute evidence for genuine bilingualism.

(6) Graeco-Arabic inscription A1 (Al-Jallad & al-Manaser 2016: 56)

a. Greek
Θαιμος Γαφαλου
Taym Gaḥfal
'Taym, son of Gaḥfal'

[13] While it is tempting to interpret *ʔn* as the first-person singular pronoun *ʔanā*, such a formula would indeed be strange in a grave epitaph. Perhaps *ʔn* is cognate with the demonstrative/presentative element *han, or perhaps it should be construed as a dative 'to, for' cognate with East Semitic *ana*.

b. Arabic
l-tm bn gḥfl
PREP-Taym son Gaḥfal
'for/by Taym, son of Gaḥfal'

The second inscription discussed by Al-Jallad & al-Manaser (2016), illustrated in (7), provides more insight into the different degrees of Arabic–Greek bilingualism.. The author carves a short text in both Greek and Old Arabic, indicating that he knew both languages but that his command of Arabic was obviously better.

(7) Graeco-Arabic inscription 2 (Al-Jallad & al-Manaser 2016: 58)

a. Arabic
l-ɣθ w tḥll ʔfwh ʕql sr
PREP-Ghawth CONJ go.PRF.3SG.M PREP protected_area Sayr
'By Ghawth and he went into the protected area of Sayr.'

b. Greek
Γαυτος ἀπῆλθεν [ε]ἰς τόν Ακελον Σαιρου
Ghawth.NOM depart.AOR.3SG PREP DEF.M.ACC.SG *ʕaql*.ACC.SG Sayr.GEN
'Ghawth, he went away to the *ʕaql* of Sayr.'

The author translates the Arabic into Greek effectively, but seems not to have known the Greek word for the culturally specific term *ʕaql*, 'a protected area of pasturage'. In this case, he simply wrote the word out in Greek: *Ακελον*.

There is evidence that some nomadic Arabic speakers did master the Greek language, as one sometimes comes across very well-composed texts in Greek, attesting to full-scale bilingualism, at least in writing (for example A2 in Al-Jallad & al-Manaser (2015). This level of bilingualism, however, must have been rare. There is no appreciable influence from Greek on the Arabic of the Safaitic inscriptions. A few loanwords are known, e.g. *qṣr* 'Caesar', *lṣṭ* 'thief', but these more likely come through Aramaic.

2.11 Arabic in eastern Arabia

The inscriptional record of eastern Arabia is relatively poor when compared to the western two-thirds of the Peninsula. Nevertheless, the extant texts point towards contact between Aramaic and the local Arabian language, called Hasaitic by scholars. This language, however, cannot be regarded as a form of Arabic, and there are no pre-Islamic attestations of Arabic from eastern Arabia yet (Al-Jallad 2018b: 260–261).

3 Grammatical features arising from contact

This section offers a contact-based explanation for two linguistic features found in Old Arabic: the definite article, and the realization of the feminine ending.

3.1 Definite article

It has long been established that the overt marking of definiteness in the Semitic languages is a relatively late innovation (Huehnergard & Rubin 2011: 260–261). All varieties of Arabic today attest some form of the definite article – most commonly variants of *ʔal* but other forms exist as well, mainly in southwest Arabia, including *am*, *an*, and *a-*, with gemination of the following consonant. In light of the comparative evidence, did Arabic innovate this feature independently or was contact with other Semitic languages involved?

The evidence suggests that the prefixed article *han- emerged in the central Levant sometime in the late second millennium BCE, after the diversification of Northwest Semitic (Tropper 2001; Gzella 2006; Pat-El 2006). It seems clear that by the early first millennium BCE, the article had spread across the southern Levant and to North Arabia, as it is found in Taymanitic, Thamudic B, and Dadanitic, as well as in the Old Arabic of the Safaitic inscriptions. In the latter case, contact with Canaanite is substantiated in the inscriptional record in the form of the Bāyir inscription (see §2.2 above).

All of these languages, including the earliest Old Arabic, took over the form of the article unchanged; that is *h-* with the assimilation of the /n/ before a consonant, the exception being Dadanitic, which preserves the /n/ before laryngeal consonants, e.g. *h-mlk* /ham-malk/ ‘the king’ vs. *hn-ʔʕly* ‘the upper’ /han-ʔaʕlay/. We cannot, however, argue for the spread of the definite article to Proto-Arabic. The original, article-less situation is attested in the inscriptions of Central Jordan stretching down to the Hismā, known as Hismaic (Graf & Zwettler 2004). These texts are in unambiguously Arabic language, but they lack the definite article. The *h*-morpheme exists, but it has a strong demonstrative force. Indeed, in a few Nabataean–Hismaic bilingual inscriptions, the definite article *ʔl* of the Nabataean component is rendered as zero in the Hismaic text (Hayajneh 2009). A minority of Safaitic inscriptions also lack the definite article (Al-Jallad 2018b), showing that it had not spread to all varieties of Arabic even as late as the turn of the Era. Thus, like Hebrew and Aramaic, the earliest linguistic stages of Arabic – and indeed Proto-Arabic – lacked a fully grammaticalized definite article. Contact with Canaanite then seems to be the likeliest explanation for the appearance of the *h*-article in Old Arabic.

While the *h-* article is the commonest form in Old Arabic, whence the *ʔal* form? The *ʔal* article appears to be a later development from the original *han* article, through two irregular sound changes: *h* > *ʔ* and *n* > *l*.[14] The former is well attested in Arabic (e.g. the causative ʔaCCaCa from haCCaCa), while the latter is not uncommon in loans (e.g. *finǧān* vs. *finǧāl* 'cup'). The *ʔal* article appears to have developed in the western dialects of Old Arabic, attested first in the Nile Delta (cf. the famous *αλιλατ al-ʔilat* 'the goddess' mentioned in Herodotus , *Histories* I: 131), and is the regular form of the article in the dialect of the Nabataeans, who were situated in ancient Edom, stretching south to the Ḥiǧāz. The *ʔal*-article is attested sporadically at Dadān in the western Ḥiǧāz as well. Based on the inscriptional record, the *ʔal*-article was a typical linguistic feature of settled, rather than nomadic groups, being attested most frequently in the Nabataean dialect, and in cities and oases like Petra and Ḥegrā. The nomads used a variety of definite article forms. It was perhaps not until the rise of Islam, and the resulting prestige given to official Arabic of the Umayyad state, that the *ʔal* article began to dominate at the expense of other forms.

3.2 The feminine ending

In most modern Arabic dialects, the feminine ending *-at is realized as *-a(h)* in all contexts except the construct state, where it retains its original form *-at*. In Classical Arabic, it is *-at* in all situations, except for in utterance-final position, where it is realized as *-ah*. The Quranic Consonantal Text resembles the situation in the modern dialects, as do the transitional Nabataeo-Arabic and sixth-century Arabic script inscriptions (Nehmé 2017). Yet, if we go back further to the first century CE, it seems that varieties of Arabic written in the Hismaic and Safaitic script never experienced the sound change *-at* > *-ah* in any position – the feminine ending is always written as ⟨t⟩. In the Arabic of the Nabataeans, however, the sound change of *-at* to *-ah* seems to have operated as early as the third century BCE (Al-Jallad 2017: §5.2.1).

The sound change *-at* > *-ah* is common in the Central Semitic languages, but the distribution can vary. In Phoenician, it applies to verbs and not nouns, while in Hebrew it applies equally to nouns and verbs (Huehnergard & Rubin 2011: 265–266). The most common Arabic distribution matches Aramaic: it applies to nouns but not verbs. I would suggest that, since this sound change is first attested in a dialect of Arabic for which we have abundant evidence of heavy contact with Aramaic, it is likely is a contact-induced change (see also van Putten, this

[14]The origins of the *al*-article are discussed in detail in (Al-Jallad forthcoming).

volume). Contact, or the lack thereof, may explain its absence in the ancient nomadic dialects, where, as we have seen above, there is little evidence for contact with Aramaic. Thus, like the *ʔal* article, the *-at* to *-ah* change would have been a typical feature of Arabic dialects of settled groups in the pre-Islamic period. In later forms of Arabic, the change spreads even to nomadic dialects, as we find it operational today across the Arabian Peninsula. Yet, the chronology of this diffusion is not quite clear. In an important study by van Putten (2017), the Dosiri dialect of Kuwait appears to preserve the archaic situation where the feminine ending is realized as *-at* in all positions.

4 Conclusion

Contact must be factored into our understanding of language change for Arabic at every attested stage. A summary of the facts above show that Arabic was in most intense contact with Aramaic, a situation that persisted for over a millennium prior to the rise of Islam, which may explain the high number of Aramaic loanwords into Arabic, and indeed some striking structural parallels, such as the distribution of the sound change *-at* > *-ah*. At the same time, there is very little evidence for contact with Sabaic (Old South Arabian), a contact situation only represented by a small number of mixed texts. This nicely matches the absence of South Arabian influence on Old Arabic and later forms of the language, with the exception of those dialects spoken in southwest Arabia.

Further reading

Al-Jallad (2018b) provides a comprehensive outline of the languages and scripts of pre-Islamic North Arabia.
Macdonald (2003) gives a description of the multilingual environment of ancient Nabataea.
Nehmé (2010) outlines the development of the Arabic script based on the newest Nabataeo-Arabic inscriptions from Northwest Arabia.
Stein (2018) gives an outline of the use of Aramaic in pre-Islamic Arabia.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	INF	infinitive
ACC	accusative	M	masculine
ACT	active	NOM	nominative
AOR	aorist	PASS	passive
BCE	before Common Era	PL	plural
c.	century	PN	proper noun
CE	Common Era	PREP	preposition
CONJ	conjunction	PRF	perfect (suffix conjugation)
DEF	definite	PTCP	participle
DEM	demonstrative	REL	relative
F	feminine	SG	singular
IMPF	imperfect (prefix conjugation)	VOC	vocative

References

Al-Jallad, Ahmad. 2015. *An outline of the grammar of Safaitic inscriptions*. Leiden: Brill.

Al-Jallad, Ahmad. 2017. The earliest stages of Arabic and its linguistic classification. In Elabbas Benmamoun & Reem Basiouney (eds.), *The Routledge handbook of Arabic linguistics*, 315–331. London: Routledge.

Al-Jallad, Ahmad. 2018a. The Arabic of Petra. In Antti Arjava, Jaakko Frösén & Jorma Kaimio (eds.), *The Petra papyri V*. Amman: American Center of Oriental Research.

Al-Jallad, Ahmad. 2018b. What is Ancient North Arabian? In Daniel Birnstiel (ed.), *Re-engaging comparative Semitic and Arabic studies*, 1–44. Wiesbaden: Harrassowitz.

Al-Jallad, Ahmad. Forthcoming. One *wāw* to rule them all: The origins and fate of wawation in Arabic. In Fred M. Donner & Rebecca Hasselbach-Andee (eds.), *Scripts and scripture*. Chicago: Oriental Institute Publications.

Al-Jallad, Ahmad & Ali al-Manaser. 2015. New epigraphica from Jordan I: A pre-Islamic Arabic inscription in Greek letters and a Greek inscription from north-eastern Jordan. *Arabian Epigraphic Notes* 1. 57–70.

Al-Jallad, Ahmad & Ali al-Manaser. 2016. New epigraphica from Jordan II: Three Safaitic–Greek partial bilingual inscriptions. *Arabian Epigraphic Notes* 2. 55–66.

Arjava, Antti, Jaakko Frösén & Jorma Kaimio (eds.). 2018. *The Petra papyri V*. Amman: American Center of Oriental Research.

Avner, Uzi, Laïla Nehmé & Christian Robin. 2013. A rock inscription mentioning Thaʿlaba, an Arab king from Ghassān. *Arabian Archaeology and Epigraphy* 24(2). 237–256.

Blau, Joshua. 1981. *The emergence and linguistic background of Judaeo-Arabic: A study of the origins of Neo-Arabic and Middle Arabic*. 3rd edn. Jerusalem: Ben-Zvi Institute.

Durand, Caroline. 2017. Banqueter pour mieux régner? A propos de quelques assemblages céramiques provenant de Pétra et du territoire Nabateen. *Syria* 74. 85–98.

Ephʿal, Israel. 1974. "Arabs" in Babylonia in the 8th century B.C. *Journal of the American Oriental Society* 94(1). 108–115.

Ephʿal, Israel. 1982. *The ancient Arabs: Nomads on the borders of the Fertile Crescent, 9th–5th centuries B.C.* Jerusalem: Magnes Press.

Fiema, Zbigniew T., Ahmad Al-Jallad, Michael C. A. Macdonald & Laïla Nehmé. 2015. Provincia Arabia: Nabataea, the emergence of Arabic as a written Language, and Graeco-Arabica. In Greg Fisher (ed.), *Arabs and empires before Islam*, 373–433. Oxford: OUP.

Graf, David F. & Michael J. Zwettler. 2004. The North Arabian "Thamudic E" inscription from Uraynibah West. *Bulletin of the American Schools of Oriental Research* 335. 53–89.

Gzella, Holger. 2006. Die Entstehung des Artikels im Semitischen: Eine "phönizische" Perspektive. *Journal of Semitic Studies* 51(1). 3–18.

Gzella, Holger. 2015. *A cultural history of Aramaic: From the beginnings to the advent of Islam*. Leiden: Brill.

Hayajneh, Hani. 2009. Ancient North Arabian–Nabataean bilingual inscriptions from southern Jordan. *Proceedings of the Seminar for Arabian Studies* 39. 203–222.

Hayajneh, Hani, Mohammad Ababneh & Fawwaz Khraysheh. 2015. Die Götter von Ammon, Moab und Edom in einer neuen frühnordarabischen Inschrift aus Südost-Jordanien. In Viktor Golinets, Hanna Jenni, Hans-Peter Mathys & Samuel Sarasin (eds.), *Neue Beiträge zur Semitistik: Fünftes Treffen der Arbeitsgemeinschaft Semitistik in der Deutschen Morgenländischen Gesellschaft vom 15.–17. Februar 2012 an der Universität Basel*, 79–105. Münster: Ugarit-Verlag.

Huehnergard, John & Aaron Rubin. 2011. Modern Standard Arabic. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 259–278. Berlin: De Gruyter Mouton.

Israel, Felice. 1995. L'onomastique arabe dans les inscriptions de Syrie et de Palestine. In Hélène Lozachmeur (ed.), *Présence arabe dans le croissant fertile avant l'Hégire*, 47–57. Paris: Editions Recherche sur les civilizations.

Kootstra, Fokelien. 2019. *The writing culture of ancient Dadān: A description and quantitative analysis of linguistic variation*. Leiden: Leiden University. (Doctoral dissertation).

Kropp, Manfred. 2017. The ʿAyn ʿAbada inscription thirty years later: A reassessment. In Ahmad Al-Jallad (ed.), *Arabic in context*, 53–74. Leiden: Brill.

Larcher, Pierre. 2010. In search of a standard: Dialect variation and new Arabic features in the oldest Arabic written documents. In Michael C. A. Macdonald (ed.), *The development of Arabic as a written language*, 103–112. Oxford: Archaeopress.

Livingstone, Alasdair. 1997. An early attestation of the Arabic definite article. *Journal of Semitic Studies* 42(2). 259–262.

Macdonald, Michael C. A. 2000. Reflections on the linguistic map of pre-Islamic Arabia. *Arabian Archaeology and Epigraphy* 11. 28–79.

Macdonald, Michael C. A. 2003. Languages, scripts, and the uses of writing among the Nabataeans. In Glenn Markoe (ed.), *Petra rediscovered: The lost city of the Nabataeans*, 37–56. New York: Harry N. Abrams/The Cincinnati Art Museum.

Macdonald, Michael C. A. 2010. Ancient Arabia and the written word. In Michael C. A. Macdonald (ed.), *The development of Arabic as a written language*, 5–27. Oxford: Archaeopress.

Negev, Avraham, J. Naveh & S. Shaked. 1986. Obodas the god. *Israel Exploration Journal* 36. 56–60.

Nehmé, Laïla. 2010. A glimpse of the development of the Nabataean script into Arabic based on old and new epigraphic material. In Michael C. A. Macdonald (ed.), *The development of Arabic as a written language*, 47–88. Oxford: Archaeopress.

Nehmé, Laïla. 2017. Aramaic or Arabic? the Nabataeo-Arabic script and the language of the inscriptions written in this script. In Ahmad Al-Jallad (ed.), *Arabic in context*, 75–89. Leiden: Brill.

Nyberg, Henrik Samuel. 1974. *A manual of Pahlavi*. Vol. 2: *Ideograms, glossary, abbreviations, index, grammatical survey, corrigenda to Part I*. Wiesbaden: Harrassowitz.

Overlaet, Bruno, Michael C. A. Macdonald & Peter Stein. 2016. An Aramaic–Hasaitic bilingual inscription from a monumental tomb at Mleiha, Sharjah, UAE. *Arabian Archaeology and Epigraphy* 27(1). 127–142.

Pat-El, Na'ama. 2006. The development of the Semitic definite article: A syntactic approach. *Journal of Semitic Studies* 51(1). 19–50.

Prioletta, Alessia & Christian Julien Robin. 2018. New research on the 'Thamudic' graffiti from the region of Ḥimā (Najrān, Saudi Arabia). In Michael C. A. Macdonald (ed.), *Languages, scripts and their uses in ancient North Arabia: Papers from the special session of the Seminar for Arabian Studies held on 5 August 2017*, 53–70. Oxford: Archaeopress.

van Putten, Marijn. 2017. The archaic feminine ending *-at* in Shammari Arabic. *Journal of Semitic Studies* 62(2). 357–369.

Rabinowitz, Isaac. 1956. Aramaic inscriptions of the fifth century B. C. E. from a North-Arab shrine in Egypt. *Journal of Near Eastern Studies* 15(1). 1–9.

Retsö, Jan. 2003. *The Arabs in antiquity: Their history from the Assyrians to the Umayyads*. London: Routledge.

Robin, Christian-Julien. 1991. La pénétration des arabes nomades au Yémen. *Revue des mondes musulmans et de la Méditerranée* 61(1). 71–88.

Souag, Lameen. 2018. Yaqṭīn *as substratum vocabulary*. http://lughat.blogspot.com/2018/06/yaqtin-as-substratum-vocabulary.html. Accessed 19/9/2019.

Stein, Peter. 2018. The role of Aramaic on the Arabian Peninsula in the second half of the first millennium BC. In Michael C. A. Macdonald (ed.), *Languages, scripts and their uses in ancient North Arabia*, 39–49. Oxford: Archaeopress.

Tropper, Josef. 2001. Die Herausbildung des bestimmten Artikels im Semitischen. *Journal of Semitic Studies* 46(1). 1–31.

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.

Yardeni, Ada. 2014. A list of the Arabic words appearing in Nabataean and Aramaic legal documents from the Judaean Desert. *Scripta Classica Israelica* 33. 301–324.

Chapter 2

Classical and Modern Standard Arabic

Marijn van Putten
University of Leiden

The highly archaic Classical Arabic language and its modern iteration Modern Standard Arabic must to a large extent be seen as highly artificial archaizing registers that are the High variety of a diglossic situation. The contact phenomena found in Classical Arabic and Modern Standard Arabic are therefore often the result of imposition. Cases of borrowing are significantly rarer, and mainly found in the lexicon sphere of the language.

1 Current state and historical development

Classical Arabic (CA) is the highly archaic variety of Arabic that, after its codification by the Arab Grammarians around the beginning of the ninth century, becomes the most dominant written register of Arabic. While forms of Middle Arabic, a style somewhat intermediate between CA and spoken dialects, gain some traction in the Middle Ages, CA remains the most important written register for official, religious and scientific purposes.

From the moment of CA's rise to dominance as a written language, the whole of the Arabic-speaking world can be thought of as having transitioned into a state of diglossia (Ferguson 1959; 1996), where CA takes up the High register and the spoken dialects the Low register.[1] Representation in writing of these spoken dialects is (almost) completely absent in the written record for much of the Middle Ages. Eventually, CA came to be largely replaced for administrative purposes by Ottoman Turkish, and at the beginning of the nineteenth century, it was functionally limited to religious domains (Glaß 2011: 836). During the nineteenth-century

[1]Diglossic situations are often seen as consisting of a high register (often called H) and a low register (L). These two are seen to be in complementary distribution, where each register is used in designated environments, where the H register takes up such domains like formal speeches and writing, while the L register is used in personal conversation, oral literature etc.

Marijn van Putten. 2020. Classical and Modern Standard Arabic. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 23–49. Berlin: Language Science Press. DOI:??

Arabic literary revival known as the *Nahḍa*, CA goes through a rather amorphous and decentralized phase of modernization, introducing many neologisms for modern technologies and concepts, and many new syntagms became part of modern writing, often calqued upon European languages. After this period, it is customary in scholarly circles to speak of Classical Arabic having transitioned into Modern Standard Arabic (MSA), despite the insistence of its authors that CA and MSA are one and the same language: *al-ʕarabiyya l-fuṣḥā* 'the most eloquent Arabic language' (Ryding 2011: 845).

2 Contact languages

Considering the significant time-depth of CA and MSA, contact languages have of course changed over time. Important sources of linguistic contact of the pre-Islamic varieties of Arabic that come to form the vocabulary for CA are Aramaic, Greek and Ethio-Semitic. While there are already some Persian loanwords in the very first sources of CA, this influence continues well into the Classical period, and ends up having a marked effect on CA and MSA alike.

2.1 Aramaic

Aramaic becomes the dominant lingua franca in much of the Achaemenid empire, and both written and spoken varieties of Aramaic continue to play an essential role all throughout Arabia, Syria and Mesopotamia right up until the dawn of Islam. As such, a not insignificant amount of vocabulary has been borrowed from Aramaic into Arabic, which shows up in Classical Arabic. Moreover, Aramaic was an important language of Christianity and Judaism, and a noticeable amount of religious vocabulary from Aramaic has entered Classical Arabic (§3.4.2). There may even be some structural influence on the phonology of pre-Classical Arabic that has made it into Arabic (§3.1).

2.2 Greek

Greek was the language of state of the Byzantine Empire, which, when not directly ruling over Arabic-speaking populations, was at least in close contact with them. This can be seen in the significant amount of Greek vocabulary that can be detected in CA. Aramaic, however, has often borrowed the same terms that we find in CA, and it is usually difficult, if not impossible, to decide whether a Greek word entered Arabic directly from Greek or through the intermediary of Aramaic (§3.4.3).

2.3 Persian

After the rise of Islam, Greek and Aramaic quickly lose the central role they once played in the region, and they do not continue to influence CA significantly in the Islamic period. Persian however, of which a number of words can already be detected in the Quran, continues to have a pronounced influence on Arabic, and many more Persian words enter CA throughout its history (§3.4.5).

2.4 Ethio-Semitic & Ancient South Arabian

It is widely recognized that some degree of influence from Ethio-Semitic can be identified within CA (§3.2.3; 3.4.1). Many of the Ethio-Semitic words that have entered into Quranic Arabic presumably arrived there through South Arabian contact after the invasion of Yemen by Christian Ethiopia in the sixth century. Also previous South Arabian contact must probably be assumed, and the divine epithet *ar-Raḥmān* is usually thought to be a borrowing from South Arabian, where it in turn is probably a borrowing from Aramaic (Jeffery 2007 [1938]: 140–141).

While Ethio-Semitic contact has been fairly well-researched, research into contact with Ancient South Arabia is still in its infancy. The exact classification of the Ancient South Arabian languages and their relation to Modern South Arabian and Ethio-Semitic is still very much under debate. A simple understanding of this highly multilingual region seems impossible. Due to the extensive contact within South Arabia and the South Arabian languages, it is not always easy to pin down the exact vector of contact between CA and these languages of South Arabia and Ethiopia (§3.4.4).

2.5 Arabic dialects

The spoken Arabic dialects, of course, have had and continue to have a noticeable influence on CA and MSA (§2.5; 3.2.1; 3.2.2; 3.3; 3.5). It seems that from the very moment CA became canonized as an official language, it was already a highly artificial register that nobody spoke in the form in which it was canonized. Especially the Ḥiǧāzi conquerors had a noticeable effect on the language – no doubt through mediation of the Quranic text. Noticeable irregularities in the treatment of the glottal stop, for example, have entered the language, and have influenced the treatment of certain morphological features (§3.2).

2.6 Ottoman Turkish

In the Ottoman period, Ottoman Turkish becomes the official language in use in the Middle East, and replaces many of the sociolinguistic functions that CA had before that. The imposition of this official language had a significant effect on the Arabic vernaculars throughout the Middle East (even outside the borders of the Ottoman Empire), but also had a noticeable impact on the vocabulary of CA, especially in the eighteenth and nineteenth centuries, which feeds into MSA (§3.4.6).

3 Contact-induced changes in Classical and Modern Standard Arabic

3.1 Phonology

Due to the highly conservative nature of CA, finding any obvious traces of contact in phonological change is very difficult. From the period in which Sibawayh describes the phonology of the *ʕarabiyya* until today, only minor changes have taken place in the phonology of CA. The most obvious example of this is the loss of the lateral realization of the *ḍād*, which in Sibawayh's description is still a lateral, while today it is generally pronounced as [dˤ]. Blau (1969: 162–163) convincingly attributes this development to influence from the modern dialects. In most modern Arabic dialects, the reflexes of *ḍ* [ɮˤ] and *ð̣* [ðˤ] merged to *ð̣* [ðˤ].[2] In sedentary dialects that lose the interdentals, this merged sound subsequently shifts to *ḍ* [dˤ]. As such, original *ð̣ and ḍ* are either both pronounced as an emphatic interdental fricative or both as an emphatic dental stop. As virtually all modern dialects, however, have lost the lateral realization of *ḍ*, the sedentary stop realization was repurposed for the realization of *ḍ*, to introduce the phonemic distinction between *ð̣* and *ḍ* in MSA.

As this is a case where the speakers influencing the phonology of the RL are SL-dominant, this change of pronunciation of the *ḍ* from a lateral to a stop realization can be seen as a form of imposition on the phonology of MSA. It should be noted, however, that the type of imposition we are dealing with in this case is of quite a different character than what is traditionally understood as imposition within the framework of Van Coetsem (1988; 2000). In this case, we see a conscious effort to introduce a phonemic distinction lost in the SL between original

[2]Not all dialects, however, see Behnstedt (2016: 16ff.).

ḍ and *ḍ̊* by using two different dialectal outcomes of the merger of these two phonemes.

Other cases of phonetic imposition on MSA from the modern dialects may especially be found in the realization of the *ǧīm*. While Sibawayh's description of the *ǧīm* was probably a palatal stop [ɟ], today the realization that seems to carry the most prestige and is generally adhered to in Quranic recitation is [ʤ]. However, here too we often find imposition of the local pronunciation of this sound in MSA. In spoken MSA of Egyptians the *ǧīm* is regularly pronounced as [g], the realization of the *ǧīm* in Egyptian Arabic. Likewise, Levantine Arabic speakers whose reflex of the *ǧīm* is [ʒ] will often use that realization when speaking MSA.

If we shift our focus to developments that began in the pre-Classical period and are continued in CA, we find that there are several phonetic developments that bear some similarity to developments of Aramaic. It has therefore, not unreasonably, been suggested that such developments are the result of contact with Aramaic.

The first of these similar phonetic developments shared between Classical Arabic and Aramaic is the shift of the semivowels *w* and *y* to *ʔ* between a preceding *ā* and a following short vowel *i* or *u*. This can be seen, for example, in the similar outcomes of the active participles of hollow roots. This similarity was already remarked upon and described by Brockelmann (1908: 138–139), e.g.:

(1) a. CA *qāwimun > *qāʔimun* 'standing'
b. Aram. *qāwim > *qāʔem* 'standing'

However, it is clear that, at least in Nabataean Arabic, this development had not yet taken place (Diem 1980: 91–93), this being a dialect that was certainly in contact with Aramaic, as most of the writing of the Nabataeans was in a form of Aramaic. As such, we may plausibly suggest that this development took place *after* the establishment of linguistic contact between Aramaic and Arabic. It is quite difficult to decide whether this development, if we are correct to interpret it as the result of contact-induced change, is the result of imposition, borrowing or convergence. We do not have a clear enough picture of the sociolinguistic relations between Aramaic and Pre-Classical Arabic to decide the type of contact situation that would have caused it. One is tempted to see it as the result of imposition simply because of the fact that phonological borrowing seems to be uncommon (Lucas 2015: 526).[3]

[3]We cannot discount the possibility of parallel development, however. Akkadian seems to have undergone an almost identical development (Huehnergard 1997: 196), where it is not likely to have been the result of contact.

As proposed by Al-Jallad (this volume), another possible case of contact induced phonological change between Aramaic and Pre-Classical Arabic is the shift of pausal *-at* to *ah*, found only in nouns and not in verbs. Huehnergard & Rubin (2011: 267–268) already suggested that this development, which cannot be due to a development in a shared ancestor, may have been the result of areal diffusion.

Whether we can really interpret the development of Aramaic as similar to that of CA, however, depends somewhat on the interpretation of the Aramaic evidence. While we can indeed see a development of the original Aramaic feminine ending **-at* that is written with ⟨-h⟩ in consonantal writing, which might suggest it has shifted to /ah/, one also finds that all other cases of word-final nominal *t* have been lost, while not leaving a consonantal *-h*, e.g.:

(2) a. *ṣalōt > *ṣlō* ⟨ṣlw⟩ 'prayer'
b. *zakūt > *zkū* ⟨zkw⟩ 'merit, victory'
c. *ešāt > *ʔešā* ⟨ʔšʔ⟩ 'fire'
d. *bayt > *bay* ⟨by⟩ 'house'

For this parallel loss of final *t* in all other environments, Beyer (1984: 96, fn. 4) prefers to interpret the ⟨-h⟩ as a *mater lectionis* for final /ā/ or /a/. In this interpretation, the development of Aramaic compared to Arabic is quite different, as in Arabic the ⟨-h⟩ is clearly consonantal, and the loss of final *t* does not happen after long vowels in Arabic:

(3) Aramaic
a. *kalbat > *kalbā* ⟨klb⟩ 'bitch' (*-at# > -a/ā*)
b. *ʔešāt > *ʔešā* ⟨ʔšʔ⟩ 'fire' (*-āt# > -ā*)

(4) Arabic
a. *kalbat > *kalbah*
b. *kalbāt > *kalbāt* remains unchanged

However, if one takes the ⟨-h⟩ of the feminine to originally represent *-at > *-ah*, and the loss of *t* in other word-final positions to be a different development, one could reasonably attribute the development in Arabic to the result of contact with Aramaic, as it is clear that in many varieties of Pre-Islamic Arabic, the *-at > *-ah* shift had not yet taken place.[4]

[4]For a discussion on the development of the *-at > *-ah* shift in Pre-Islamic Arabic see Al-Jallad (2017: 157–158).

3.2 Morphology

3.2.1 Imposition of the taCCiʔah stem II verbal noun for glottal-stop-final verbs

A well-known feature of Ḥiğāzī Arabic in the early Islamic period, and a feature that is found in many of the modern dialects is the (almost) complete loss of the glottal stop (Rabin 1951: 130–131; van Putten 2018). This loss has usually caused glottal-stop-final roots to be reanalyzed as final-weak verbs, e.g. Cairene *ʔara, ʔarēt* 'he read, I read' (< *qaraʔa, *qaraʔtu).

A typical feature of final-weak verbal noun formations in CA is their formation of the verbal noun of stem II verbs. Sound verbs form verbal nouns using the pattern taCCīC, e.g. *taslīm* 'greeting' from *sallama* 'to greet'. Final-weak verbs, however, regularly use the pattern taCCiyah instead (Fischer 2002: 44), for example, *tasmiyah* 'naming' from *sammā* 'to name'.[5]

In CA, the *ʔ* generally functions as a regular consonant. Thus a verb like *qaraʔa/yaqraʔu* 'to read, recite' does not differ significantly in its behavior from any other triconsonantal verb such as *fataḥa/yaftaḥu* 'to open'.

However, verbs with *ʔ* as final root consonants unexpectedly frequently side with the final-weak verbs when it comes to the verbal noun of stem II verbs (Fischer 2002: 128). For example, *hannaʔa/yuhannaʔu* 'to congratulate' does not have the expected verbal noun **tahnīʔ, but instead *tahniʔah* 'congratulation'. Other examples are:

(5) a. *nabbaʔa* v.n. *tanbiʔah* (besides *tanbīʔ*) 'to inform'
b. *barraʔa* v.n. *tabriʔah* 'to acquit'
c. *hayyaʔa* v.n. *tahyiʔah* (besides *tahyīʔ*) 'to make ready'
d. *naššaʔa*, v.n. *tanšiʔah* (besides *tanšīʔ*) 'to raise (a child)'

Some other verbs with the same pattern do have the expected CA form such as *baṭṭaʔa* v.n. *tabṭīʔ* 'to delay'.

This behaviour can plausibly be attributed to the fact that in many (if not most) spoken varieties of Arabic, from early on the final-glottal-stop verbs had already merged completely with the final-weak verbs, and as such a verb like *hannaʔa* had come to be pronounced as *hannā*, and thus reanalyzed as a final-weak verb.

[5]This is an ancient idiosyncrasy of final-weak verbs. While the taCCīC formation is not a regular formation in other Semitic languages, when it does occur, the final-weak verbs have a feminine ending, e.g. Hebrew *tarmi-ṯ* 'betrayal', *toḏå* 'praise' (< *tawdiy-ah or *tawday-ah), see Brockelmann (1908: 385–387).

Like original final-weak verbs, their regular verbal noun formation would be like *tahniyah*. When verbs of this type were employed in CA, the weak root consonant *y* was replaced with the etymological glottal stop *ʔ*, rather than completely converting the verbal noun to the regular pattern. This is a clear example of the imposition of a morphological pattern onto CA grammar by speakers of Arabic dialects.

3.2.2 Imposition of the ʔaCCiyāʔ broken plural pattern

A similar case of imposition, where the morphological categories of glottal-stop-final roots behave in the grammar as if they are final-weak, may be found in the broken-plural formation of CaCīʔ nouns and adjectives. The broken-plural formation most generally used for final-weak adjectives with the pattern CaCiyy (< *CaCīy) is ʔaCCiyāʔ. For example, *ɣaniyy* pl. *ʔaɣniyāʔ* 'rich', *waliyy* pl. *ʔawliyāʔ* 'close associate', *daʕiyy* pl. *ʔadʕiyāʔ* 'bastard', *sawiyy* pl. *ʔaswiyāʔ* 'correct', *ḫaliyy* pl. *ʔaḫliyāʔ* 'free'.

For sound nouns of this type, it is much more typical to use the plural formations CiCāC (*kabīr* pl. *kibār* 'big') or CuCaCāʔ (*faqīr* pl. *fuqarāʔ* 'poor'), although there are a couple of sound nouns that do use this plural, such as *qarīb* pl. *ʔaqribāʔ* 'relative' and *ṣadīq* pl. *ʔaṣdiqāʔ* 'friend' (Ratcliffe 1998: 106–107).[6]

CaCīC formations where the last root consonant is *ʔ*, however, behave in rather unexpected ways in the CA, usually following the pattern of final-weak nouns, often even replacing the final *ʔ* with *y*, for example: *barīʔ* pl. *ʔabriyāʔ* 'free', *radīʔ* pl. *ʔardiyāʔ* 'bad'. These nouns have plurals that are proper not to the Classical form they have, but rather to the colloquial form without *ʔ*, i.e. *bariyy*, *radiyy*. Once again this can be seen as a clear case of imposition of the colloquial Arabic forms onto the classical language.[7]

[6]The pattern (with metathesis) is also regular for geminated CaCīC adjectives, e.g. *šadīd* pl. *ʔašiddāʔ* 'severe'.

[7]These two cases of imposition of glottal stop-less morphology onto CA are two of the more clear and systematic cases, but close observations of CA morphology reveals many more of these somewhat more isolated cases, e.g. *ḫaṭīʔah* 'sin' with a plural *ḫaṭāyā* which is rather the expected plural of *ḫaṭiyyah*; *bariyyah* pl. *barāyā* 'creature' which is a derivation from *baraʔa* 'to create'; *ðurriyyah*, *ðirriyyah* pl. *ðarāriyy* 'progeny, offspring', derived from *ðaraʔa* 'to sow, seed'. Another example of irregular treatment of *ʔ* that is presumably the result of impositition is found in verbal nouns of stem VI verbs, and *mafāʕil* plurals of hollow roots, which modern textbooks say should not have a *ʔ* despite having the environment that is expected to undergo the shift *āwu/i, āyi* > *āʔu/i, āʔi* as discussed in §3.1. The lexicographical tradition and Quranic reading traditions often record disagreements on the application of the *hamzah* in such cases. For example, we find both *tanāwuš* and *tanāʔuš* 'reaching one another', and *maʕāyiš* and *maʕāʔiš* 'ways of living'.

3.2.3 Borrowing of the broken plural pattern CaCāCiCah

CA, like the modern Arabic dialects, is well-known for its broken-plural patterns. This is a feature it shares especially with Ancient South Arabian (Stein 2011: 1050–1051), Modern South Arabian Languages (Simeone-Senelle 2011: 1085) and Ethio-Semitic (Weninger 2011b: 1132). The use of broken plurals has caused somewhat of a controversy in the subgrouping of the Semitic language family. Scholars that consider broken plurals a shared retention do not view their presence as important for grouping Arabic, Ancient South Arabian, the Modern South Arabian Languages and Ethio-Semitic together (Huehnergard 2005: 159–160); while those that consider it an innovation in a subset of Semitic languages see this as a strong indication that these languages should be grouped together into a South Semitic branch (e.g. Ratcliffe 1998).

While most scholars today seem to agree that the broken-plural system is a shared retention (Weninger 2011a: 1116), it seems clear that the retention of a highly productive broken-plural system is to be considered an areal feature that clusters around South Arabia and the Horn of Africa. Classical Arabic partakes in this areal feature.

A possible case of influence from Ancient South Arabian (and/or Ethio-Semitic) into Arabic is the introduction of the CaCāCiCah plural formation. In the South Arabian languages,[8] the equivalent plural formation CaCāCiCt is extremely productive, and numerous words with four consonants form their plural in this way. For example in Sabaic, mCCCt is the regular plural formation to mCCC nouns of location, e.g. *mḥfd* pl. *mḥfdt* 'tower' (Beeston 1962: 34). It is likewise common in Gəʕəz, e.g. *tänbäl* pl. *tänabəlt* 'ambassador' (Dillmann 2005 [1907]: 309), and occurs occasionally in Modern South Arabian, e.g. Mehri *məlēk* pl. *məlaykət* 'angel' (Rubin 2010: 68).

While this pattern exists in CA, it is much rarer than the other broken plural formations of four consonantal forms, i.e. CaCāCiC and CaCāCīC. In the Quran, *malak* pl. *malāʔikah* 'angel' is the only plural with this pattern. This noun is widely recognized as being a loanword from Gəʕəz *malʔak, malāʔəkt* (Jeffery 2007 [1938]: 269), in part on the basis that it shares this plural formation: the word seems to have been borrowed together with its plural formation. Considering the rarity of this pattern in Arabic and how common it is in South Arabian, it seems possible that the pattern was introduced into Arabic through South Arabian contact. However, the absence of other clearly identifiable South Arabian loanwords with this plural pattern makes it quite difficult to make a strong case

[8] South Arabian is used here as a purely geographical descriptive term, not one of classification.

for this identification.

Another possible word of South Arabian origin with this plural pattern is *tubbaʕ* pl. *tabābiʕah* 'a Yemenite king', but evidence that this word is indeed of Ancient South Arabian origin is missing. The word does not occur as a separate word in Ancient South Arabian, and instead is only the first part of several Ancient South Arabian theophoric names such as *tbʕkrb*, *tbʔʕl*. Such names should probably be understood as being related to the root *√tbʕ* which, like in Arabic may have had the meaning 'following', so such names likely mean 'follower of the deity KRB' and 'follower of the deity ʔL'. Such names being associated with Yemenite kings may have led to the Arabic meaning of *tubbaʕ* as 'Yemenite king', but in Ancient South Arabian itself it does not seem to have carried a meaning of this kind.

All in all, the evidence for this really being a pattern that is the result of South Arabian influence is rather slim, although the rarity of the pattern in CA does make it look unusual. If the interpretation of this plural pattern as being a borrowing from South Arabian is correct, it seems that some South Arabian nouns were borrowed along with their respective plural. This would be a case of morphological borrowing rather than the more common type of morphological influence through imposition.[9]

Note that this plural pattern has become the productive plural pattern for quadriconsonantal loanwords regardless of them being of South Arabian origin or elsewhere, e.g. *biṭrīq* pl. *baṭāriqah* 'patrician' (< Latin *patricius*), *ʔusquf* pl. *ʔasāqifah* 'bishop' (< Greek *epískopos*), *ʔustāð* pl. *ʔasātiðah* 'master' (< Middle Persian *ōstād*), *tilmīð* pl. *talāmiðah* 'student' (< Aramaic *talmīḏ*).

3.3 Syntax

Due to CA being the High register in a diglossic situation for centuries, we should presumably consider the majority of the written material produced in this language to be written exclusively by non-native speakers. Moreover, a large proportion of its writers all throughout its written history must have been speakers not only of Arabic vernaculars but also of entirely different languages such as Persian and Turkish. It seems highly unlikely that such a multilingual background of authors of Classical Arabic would have been completely without effect on the syntax of the language; however, as it is difficult to decide from what moment

[9]This can be seen as a type of "Parallel System Borrowing" similar to that which we find in Berber languages. Berber languages, like Arabic, have apophonic plurals; but Arabic nouns are simply borrowed along with their own Arabic broken plurals (Kossmann 2010).

onward we can speak of true diglossia, and what the syntax was like before that period, it seems that tracing such influences has not yet been undertaken in detail.

There is, however, promising research being done on influence on MSA syntax from the speakers of modern Arabic dialects. Wilmsen (2010) convincingly describes one such point of influence in a paper on the treatment of object pronouns in Egyptian and Levantine newspapers.

Wilmsen (2010: 104) shows that, in the case of ditransitive verbs, Egyptian and Levantine have a different natural word order. In Egyptian Arabic, the direct object must precede the indirect object as in (6), while in Levantine Arabic the indirect object preceding the direct object is preferred, as shown in (7):

(6) Egyptian
rabbi-na yḫalli-**hū-l-ak**
Lord-1PL keep.IMPF.3SG.M-3SG.M-DAT-2SG.M
'Our lord keep him for you.'

(7) Levantine
aḷḷa yḫallī-l-ak iyyā
God keep.IMPF.3SG.M-DAT-2SG.M ACC.3SG.M
'God keep him for you.'

Wilmsen argues that the following two variant sentences in a Reuters news story written in MSA, the original in (8), likely written by an Egyptian, and the slightly altered version in (9), which appeared in a Lebanese newspaper, show exactly this difference of word order found in the respective spoken dialects:

(8) MSA (Egyptian)
al-ʔawrāq-i llatī **sallamat-hā** **la-hu** ʔarmalat-u
DEF-papers-OBL REL.SG.F give.PRF.3SG.F-3SG.F DAT-3SG.M. widow-NOM
ʕabdi l-wahhāb
personal_name
'the papers, which Abdel Wahhab's widow had **given him**'

(9) MSA (Lebanese)
al-ʔawrāq-i llatī **sallamat-hu** **ʔiyyā-hā** ʔarmalat-u
DEF-papers-OBL REL.SG.F give.PRF.3SG.F-3SG.M ACC-3SG.F widow-NOM
ʕabdi l-wahhāb
personal_name
'the papers, which Abdel Wahhab's widow had **given him**'

Wilmsen (2010: 114–115) goes on to examine three newspapers (the London-based, largely Lebanese, *al-Ḥayāt* of the years 1996–1997; the Syrian *al-Θawra* of the year 2005 and the Egyptian *al-ʔAhrām*), and shows that with the two most common verbs in the corpus with such argument structure (*manaḥa* 'to grant' and *ʔaʕṭā* 'to give'), the trend is consistently in favor of the pattern found. The recipient–theme order is overwhelmingly favoured in the Levantine newspapers, while the theme–recipient order is clearly favoured by the Egyptian newspaper. The results are reproduced in Tables 1 and 2.

Table 1: Occurences of theme–recipient and recipient–theme order with *manaḥa* 'to grant'

Database	theme–recipient	recipient–theme
al-Ḥayāt 96	29	56
al-Ḥayāt 97	27	52
al-Θawra	27	66
al-ʔAhrām	44	8

Table 2: Occurrences of theme–recipient and recipient–theme order with *ʔaʕṭā* 'to give'

Database	theme–recipient	recipient–theme
al-Ḥayāt 96	11	23
al-Ḥayāt 97	8	22
al-Θawra	9	38
al-ʔAhrām	33	2

From this data it is clear that the dialectal background of the author of an MSA text can indeed play a role in how its syntax is constructed, despite both resulting

sentences being grammatically acceptable in CA/MSA.[10]

This (and any contact phenomenon in MSA–dialect diglossia) should be seen as a case of imposition, where the dialect SL, in which the speakers/writers are dominant, has influenced the MSA RL.

It stands to reason that such syntactic research could be undertaken with CA works as well. Taking into account the biographies of authors, it might be possible to find similar imposition effects that can be connected to different dialects and languages in former times. To my knowledge, however, this work has yet to be undertaken.

3.4 Lexicon

In terms of lexicon, Jeffery's indispensable (2007 [1938]) study of the foreign vocabulary in the Quran allows us to examine some of the important sources of lexical influence on Pre-Classical Arabic. Influence from Greek, Aramaic, Gəʕəz and Persian are all readily recognizable.

3.4.1 Gəʕəz

Nöldeke (1910) is still one of the most complete and important discussions of Gəʕəz loanwords in CA. Both Gəʕəz and Arabic display a significant amount of religious vocabulary that is borrowed from Aramaic. It is quite often impossible to tell whether Arabic borrowed the word from Gəʕəz or from Aramaic. Such examples are *ṭāɣūt* 'idol', Gz. *ṭaʕot*, Aram. *ṭāʕū* 'error, idol' (Nöldeke 1910: 48); *tābūt* 'ark; chest', Gz. *tabot* 'ark of Noah, ark of the covenant'; Aram. *tēḇō* 'chest; ark' (Nöldeke 1910: 49).

There is religious vocabulary that is unambiguously borrowed from Gəʕəz, e.g. *ḥawāriyyūn* 'disciples' < Gz. *ḥäwarəya* 'apostle' and *muṣḥaf* 'book (esp. Quran)' < Gz. *mäṣḥäf* 'scripture', but there is also religious vocabulary borrowed unambiguously from Aramaic, e.g. *zakāt* 'alms' < Aram. *zāḵū* 'merit, victory'; *sifr* 'large book' < Aram. *sp̄ar, sep̄rā*. It is therefore just as likely that Arabic would have borrowed such Aramaic loanwords via Gəʕəz, as directly from Aramaic.

Some religious vocabulary from Aramaic and Hebrew can be shown to have arrived in Arabic through contact with Gəʕəz, since these words have undergone specific phonetic developments shared between CA and Gəʕəz but absent in the source language. As these often involve core religious vocabulary, and the Christian Axumite kingdom was established centuries before Islam, it seems

[10] Other works that discuss clear cases of country-specific language use of MSA include Ibrahim (2009), Parkinson (2003), Parkinson (2007) and Parkinson & Ibrahim (1999).

reasonable to assume such words to be borrowings from Gəʕəz into CA, e.g. CA *ǧahannam* 'hell' < Gz. *gähännäm* (but Hebrew *gehinnom* and Syriac *gehannā*) and CA *šayṭān* 'Satan' < Gz. *śäyṭan* (but Hebrew *śåṭån* and Syriac *sāṭānā*).[11]

3.4.2 Aramaic

As already remarked upon by Retsö (2011), Aramaic loanwords in Classical Arabic often have an extremely archaic character. The Aramaic variety that influenced Quranic and pre-Classical Arabic has not undergone the famous *bəḡaḏkəp̄aṯ* lenition of post-vocalic simple stops, nor had it lost short vowels in open syllables. This necessarily means that the form of Aramaic that influenced Quranic and Classical Arabic, even the religious vocabulary, cannot be Syriac, which almost certainly underwent both shifts before becoming a dominant religious language. The *bəḡaḏkəp̄aṯ* spirantization can be dated between the first and third centuries CE, and the syncope of short vowels in open syllables takes place sometime in the middle of the third century (Gzella 2015: 41–42). However, Classical Syriac itself, as an important vehicular language of Christianity, only emerges in the fourth century CE, well after these developments had taken place (Gzella 2015: 259).

Had *bəḡaḏkəp̄aṯ* taken place, we would expect Syr. *ḡ*, *ḏ*, *ḵ*, and *ṯ* to be borrowed with their phonetic equivalents in CA: *ɣ*, *ð*, *ḫ* and *θ* respectively.[12] This, however, is not the case; instead these consonants are consistently borrowed with the stop equivalents *ǧ*, *d*, *k* and *t*, and without the loss of vowels in open syllables, clearly showing that these Aramaic loanwords predate the phonetic developments in Classical Syriac.

(10) a. *malakūt* 'kingdom', Syr. *malḵūṯ-ā* 'kingdom' < *malakūt-ā
b. *malik* 'king', Syr. *mleḵ* 'king' < *malik[13]
c. *masǧid* 'place of worship, mosque', Syr. *masgeḏ-ā* 'place of worship' < *masgid-ā

[11] Leslau (1990) often reverses the directionality of such borrowings, though without an explanation as to why he thinks a borrowing from CA into Gəʕəz is more likely.

[12] Retsö (2011) suggests that *ḇ* could also be borrowed as *w*. This might be true, but at least the phonetic match in this case is not perfect.

[13] This word is not recognized as an Aramaic loanword by Jeffery (2007: 270), but it likely is. All the Semitic cognates of this noun are derived from a form *malk, which should have been reflected in Classical Arabic as *malk*. However, we find it with an extra vowel between the last two root consonants. This can be best understood as the epenthetic vowel insertion as it is attested in Aramaic which was then subsequently borrowed with this epenthesis into Arabic. I thank Ahmad Al-Jallad for pointing this out to me.

Even the proper names of Biblical figures have a markedly un-Syriac form.

(11) a. *zakariyā, zakariyāʔ*, Syr. *Zḵaryā* < *zakaryā
b. *mīkāʔīl, mīkāʔil*,[14] Syr. *mīḵāʔel* < *mīkāʔēl

In other words, far from Syriac being "undoubtedly the most copious source of Qurʾānic borrowings" (Jeffery 2007 [1938]: 19), the Aramaic vocabulary in the Quran seems to not be Syriac at all.[15] Any isogloss that would allow us to identify it as such is conspicuously absent. This has important historical implications, as the presence of supposed Syriac religious vocabulary in the Quran is viewed as an important indication that Syriac Christian thought had a pronounced influence on early Islam (e.g. Mingana 1927: 82–90; Jeffery 2007 [1938]: 19–22).[16] While this is of course still a possibility, this has to be reconciled with the fact that the majority of clearly monotheistic religious vocabulary was already borrowed from a form of Aramaic before the rise of Syriac as a major religious language.

This does not mean that CA is completely devoid of Aramaic loanwords that have undergone the lenition of the consonants, and several post-Quranic loanwords have been borrowed from a variety which, like Syriac, had lenited its stops, e.g.:

(12) a. *tilmīð* 'student' < Syr. *talmīḏā* (Fraenkel 1886: 254)
b. *tūθ, tūt* 'mulberry' < Syr. *tūṯā* (Fraenkel 1886: 140)
c. *ḥiltīθ, ḥiltīt* '*asa foetida*' < Syr. *ḥeltīṯā* (Fraenkel 1886: 140)
d. *kāmaḥ, kāmiḥ* 'vinegar sauce' < Syr. *kāmḵā* (Fraenkel 1886: 288)
e. *karrāθ, kurrāθ* 'leek' < Syr. *karrāṯā* (Fraenkel 1886: 144)

It is interesting to note that Aramaic loanwords in Gəʕəz reflect a similar archaicity, in those cases where this is detectable. The expected lenited *ḵ* is not rep-

[14] Most readers of the Quran read either *mīkāʔīl* or *mīkāʔil*, only the most dominant tradition today, that of Ḥafṣ, reads it in the highly unusual form *mīkāl* (Ibn Muǧāhid no date: 166).

[15] Note that Jeffery (2007 [1938]: 19) explicitly states that by Syriac he means any form of Christian Aramaic, so most notably besides Syriac also Christian Palestinian Aramaic. However, this caveat hardly solves the chronological problem, the latter rises to prominence even later.

[16] Even if we were to accept the possibility that the dating of the lenition and syncope is somehow off by several centuries, the suggestion that "it is possible that certain of the Syriac words we find in the Qurʾān were introduced by Muḥammad himself" (Jeffery 2007 [1938]: 22) must certainly be rejected. In the grammatical works of Jacob of Edessa (640–708 CE) we have an unambiguous description of the lenition of the consonants (Holger Gzella p.c.). It seems highly unlikely that a wholesale lenition took place in only a few decades between the composition of the Quran and the time of his writings.

resented with Gəʕəz *ḫ* but with *k*, and short vowels in open syllables are retained. This might suggest that, when looking for religious influences on Islam, we should rather shift our focus to the south, where during the centuries before Islam both Judaism and Christianity were introduced, presumably through the vector of Gəʕəz. Some examples of such similarly archaic Aramaic loanwords in Gəʕəz are cited by Nöldeke (1910: 31–46), e.g.:

(13) Gəʕəz

a. *mälʔäk* 'angel', cf. CA *malak*, Syr. *malʔaḵ-ā* < *malʔak-ā
b. *mäläkot* 'kingdom', cf. CA *malakūt*, Syr. *malḵūṯ-ā* < *malakūt-ā
c. *ḥämelät* 'mantle, headcloth', Syr. *ḥmīlṯ-ā* < *ḥamīlat-ā
d. *näbīy* 'prophet', cf. CA *nabiyy*, Syr. *nḇīʔ-ā* < *nabīʔ-ā
e. *mäsīḥ* 'Messiah', cf. CA *al-masīḥ*, Syr. *mšīḥ-ā* < *masīḥ-ā
f. *siʔol* 'hell', cf. Syr. *siwūl* < *siʔūl (cf. Hebr. *səʔol*)
g. *ʔärämi, ʔärämāwi, ʔäräma*y 'heathen', cf. Syr. *ʔarmāy-ā* < **ʔ*aramāy-ā
h. *mänarät, mänarat* 'candlestick', cf. CA *manārah*, Syr. *mnārṯ-ā* < *manārat-ā

As of yet, there is not a clear historical scenario that helps us better understand how both CA and Gəʕəz, and, from the scanty information that we currently have, also Ancient South Arabian, ended up with similarly archaic forms of Aramaic. This seems to suggest an as yet unattested, very archaic form of Aramaic in South Arabia. Alternatively, the syncope and lenition so well-known in Syriac may have had a much less broad distribution across the written Aramaic dialects than previously thought.[17]

3.4.3 Greek (and Latin)

Besides this noticeable cluster of Aramaic and Gəʕəz words, there are of course also Greek loanwords in CA, generally in the semantic fields of economy and administration. Very often Aramaic likewise has these words, and it is usually not possible to decide whether Arabic borrowed the word from Aramaic or directly from Greek. The former direction is presumably more likely considering the broad presence of Aramaic as a lingua franca. Some examples are e.g. *dīnār* 'dinar', Aram. *dēnār*, Gk. *dēnárion*, Lat. *denarius*; *zawǧ* 'spouse, pair', Aram. *zōḡ* 'id.', Gk. *zeûgos* 'yoke', *ṣirāṭ* 'way', Aram. *ʔeṣṭrāṭ* 'street', Gk. *stráta*, Lat. *(via)*

[17] I hope to discuss the questions raised by these loanwords in a future publication.

strata; *qirṭās* 'parchment, papyrus', Aram. *qarṭīs*, Gk. *kʰártēs*; *qaṣr* 'castle', Aram. *qaṣrā*, Gk. *kástron*, Lat. *castrum*; *qalam* 'reed-pen', Gk. *kálamos* 'reed-pen'.[18]

A new influx of mostly philosophical and scientific Greek vocabulary entered CA during the early Abbasid period (mid 8th–10th centuries), at the time of the Graeco-Arabic translation movement (Gutas 1998). Once again, these words seem to have entered the language through Syriac (Gutas 2011). From this translation movement, we have words such as *ǧins* 'genus' < Syr. *gensā* < Gk. *génos*; *faylasūf* 'philosopher' < Syr. *pīlōsōpā* < Gk. *pʰilósopʰos*; *kīmyāʔ* 'alchemy' < Syr. *kīmīyā* < Gk. *kʰēmeía* and *ʔistāðiyā* 'stadium'[19] < Syr. *estaḏyā* < Gk. *stádion*.

3.4.4 Ancient South Arabian

It is often difficult to establish from which of the South Arabian languages a certain word may come. As Ancient South Arabian retained all the Proto-Semitic consonants, a borrowing from Ancient South Arabian or an inheritance from Proto-Semitic is often difficult to keep apart in CA. While Jeffery (2007 [1938]: 305) identifies a fair number of possible words of South Arabian origin, hardly ever does it seem the only possibility. Another issue with identifying South Arabian loanwords is that we have very scanty knowledge of its vocabulary or its linguistic developments. As a result, Ancient South Arabian identifications can be quite difficult to substantiate.

In recent years several lexical studies have come out that try to draw connections between Ancient South Arabian and Arabic vocabulary, but this is often based on certain semantic extensions or uses of words as described in CA dictionaries. While these observations may end up being correct, it is somewhat difficult to evaluate whether we are truly dealing with borrowings in these cases, and the extremely limited knowledge that we have of the vowel system of the different Ancient South Arabian languages makes it difficult to evaluate this in detail. Several interesting suggestions are given by Weninger (2009), Hayajneh (2011), Elmaz (2014) and Elmaz (2016).

To illustrate the difficulties we run into when trying to identify Ancient South Arabian borrowings in Arabic, let us examine the word *tārīḫ* pl. *tawārīḫ* 'date'.

[18] Nöldeke (1910: 50 argues that the CA *qalam* must come from Greek through Gz. *qäläm*. While this is possible, there is nothing about this word that requires us to assume this directionality, nor is it particularly unlikely that CA and Gəʕəz independently borrowed this word without its Greek ending *-os*.

[19] Note here the apparent application of the Syriac lenition being borrowed as such in Arabic, unlike earlier loans. But it may also be possible that the lenition is part of the Greek lenition of the *delta* instead, as we see it today in modern Greek.

From the perspective of CA morphology, *tārīḫ* could only be a hypocorrect form of *taʔrīḫ* – which is indeed an attested biform of *tārīḫ*. The existence of the plural *tawārīḫ* rather than *taʔārīḫ*, however, seems to suggest that *taʔrīḫ* is rather a hypercorrect insertion of *hamzah* from an original form *tārīḫ*, which certainly looks foreign in its formation.

Both Hebbo (1984: 27) and Weninger (2009: 399) have suggested that this word is to be connected with the the widespread Semitic root *√wrḫ*, related to 'month' or 'moon' (cf. Hebrew *yɛraḥ* < **warḫ* 'month') which exists in Ancient South Arabian but not in Classical Arabic.[20] The verb *ʔarraḫa* 'to date' would then reasonably be taken as a backformation from *tārīḫ*.

However, this explanation still leaves us with many problems. There is perhaps some reason to suppose that in Ancient South Arabian **aw* would have collapsed to an unknown monophthong (Early Sabaic *ywm* 'day'; Late Sabaic *ym*). This this might explain why the word is *tārīḫ* and not ***tawrīḫ*, but *tārīḫ* is not actually attested in Ancient South Arabian. So while the suggestion is certainly possible, it seems that another of the many non-Arabic Ancient Northern Arabian epigraphic languages could likewise have been an origin. Barring further discoveries, many such proposed etymologies remain highly speculative, and drastically simplify the rather complex multilingual situation of pre-Islamic Arabia, where many other sources besides Ancient South Arabian remain possible (Al-Jallad 2018).

3.4.5 Persian

Whereas with the advent of Islam the influence of Aramaic, Greek and Gəʕəz on CA quickly diminished and disappeared, the influence of Persian actually increased. While the Quran already contains a sizeable number of Persian borrowings, this only increases in the following centuries.

Some clear Persian borrowings in the Quran include: *ʔistabraq* 'silk brocade', cf. New Persian *istabra* (Eilers 1962: 204); *numruq* 'cushion' < Middle Persian *namrag*; *kanz* 'treasure' < Middle Persian *ganz/ganǧ* 'treasury' (Eilers 1962: 206). Outside of the Quran many other Persian words may be found in Arabic, e.g. *dīwān* 'archive, collected writings' < Early New Persian *dīwān* (Eilers 1962: 223), *banafsaǧ*, *manafsaǧ* 'violet' < Middle Persian *banafš* (Eilers 1971: 596); *barnāmaǧ* 'program' < Middle Persian *bārnāmag* (Eilers 1962: 217-218); *wazīr* 'minister' < Middle Persian *wizīr* (Eilers 1962: 207).[21]

[20]Note, however, that the root *√wrḫ* 'month' is attested unambiguously in the singular, dual (*wrḫn*) and plural (*ʾrḫ*) in the Old Arabic corpus of Safaitic inscriptions (Al-Jallad 2015: 353).

[21]I thank Chams Bernard for updating the transcription of the Middle Persian forms.

3.4.6 Ottoman Turkish

The influence of Ottoman Turkish on MSA is significantly less than on the modern Arabic dialects, largely due to linguistic purism (Procházka 2011). Words that have entered MSA are words related to administration, technology and food, but also several other origins are found. For example: *damya* 'stamp' < *damga*; *ğumruk* 'customs' < *gümrük* (ultimately from Latin *commercium*); *bāšā* 'pasha' < *paşa*; *bābūr* < *vapur* 'steam ship' (ultimately from French [*bateau à*] *vapeur*); *quṣāğ* 'pliers' < *kıskaç*; *balṭa* 'axe' < *balta*; *šāwurma, šāwirma* 'lamb, etc., roasted on a spit' < *çevirme*; *qāwurma, qāwirma* 'fried meat' < *kavurma*; *kufta* 'meatballs' < *köfte*.

Of some interest is the *-ci* suffix that denotes professions and characterizations in Turkish. This suffix has developed some amount of productivity in modern dialects (especially in Iraq, Syria and Egypt), where it may even be suffixed to nouns of non-Turkish origin. In MSA the suffix is attested not infrequently, although it would probably go too far to say that it is productive. Some examples are *nawbatğī* 'on duty; command of the guard' < *nawba* 'shift, rotatation' + *-ci*; *qahwağī* 'coffeehouse owner' < *qahwa* 'coffee' + *-ci*; *xurdağī* 'dealer in miscellaneous smallwares' < *hordaci* 'id.'; *balṭağī* 'sapper, pioneer' < *baltaci* 'sapper'; *būyāğī* 'painter, bootblack' < *boyaci* 'painter'

3.4.7 Influence of Standard Average European

A rather different, but nevertheless important factor of language contact for MSA, especially in the journalistic style, was described by Blau (1969). Blau argues that, under influence of what he dubs "Standard Average European" (SAE; cf. Whorf 1956), MSA (as well as Modern Hebrew) has taken on large amount of vocabulary,[22] phraseology and syntax similar to the journalistic language use of European languages, which in turn share so many similarities amongst each other due to mutual influence that Blau deemed it appropriate to refer to them combined as "Standard Average European", while the actual languages of influence could be quite different in different countries (e.g. Russian and Yiddish for Modern Hebrew; English for Egyptian MSA, French for Lebanese, Moroccan, Tunisian and Algerian MSA).[23] Examples of such influence takes up over a hundred pages in Blau's pioneering work.

[22]For further discussion of the development of Modern Standard Arabic technical vocabulary see Dichy (2011) and Jacquart (1994).

[23]The influence of French in terms of borrowings and adaptations is especially salient in literary Arabic as used in the Maghreb. Kropftisch (1977) is an excellent study on this topic.

Blau identifies examples of lexical expansion of existing words to include lexical associations present in SAE, e.g. *saṭḥī* 'flat' is extended in meaning towards 'superficial' due to influence of, e.g. French *superficiel* and German *oberflächlich* (Blau 1969: 65); *ǧaww* 'air, atmosphere' comes to be used in a metaphorical sense in the same way English uses 'atmosphere', e.g. *ǧawwu s-siyāsati mukahrabun* 'the political atmosphere is electrified' (Blau 1969: 69).

Even whole phrases may show up as loan translations, such as MSA *ʔanqaða l-mawqifa* 'to save the situation', cf. French *sauver la situation*, German *die Situation retten*; MSA *qatala l-waqta* 'to kill time', cf. French *tuer le temps,* German *die Zeit totschlagen* (Blau 1969: 76). Even such highly specific metaphorical expressions as 'to miss the train', in the meaning of missing an opportunity, appears in MSA *ʔasriʕ wa-ʔillā fātaka l-qiṭāru* 'hurry, otherwise you will miss the train' (Blau 1969: 101).

Such linguistic influence, of course, does not lend itself particularly well to be classified within the framework of Van Coetsem (1988; 2000), as the writers of MSA in these cases are dominant in neither the source language(s) nor the recipient language, a situation which is a rather unique result of the Arabic diglossia in combination with the influence of foreign journalistic styles that have transformed the way in which MSA is written.

3.5 Influence of the early Islamic vernaculars

While, as a general rule, Classical Arabic retains its archaic features, such as the retention of glottal stop in all positions and the lack of vowel harmony and syncope, we occasionally find single lexical items which optionally allow innovative forms which presumably stem from spoken vernaculars before the standardization of the classical language. This tends to be visible especially for words that have lost the glottal stop, a feature usually attributed to the Ḥiǧāzī variety of the early Islamic period. For example, CA has *nabiyy* 'prophet', *nubuwwah* 'prophethood' from the root *√nbʔ*;[24] likewise *bariyyah* 'creature' from the root *√brʔ*.[25]

The likely loss of postconsonantal *ʔ* in Ḥiǧāzī Arabic has influenced the way the verb *raʔā* 'to see' (*√rʔy*) is conjugated. Its imperfect irregularly loses the *ʔ*: *yarā* 'he sees'. Similarly the verb *saʔala* 'to see' (*√sʔl*) has two different imperatives, either the regular *isʔal* or the Ḥiǧāzī *sal* (< *sʔal). The imperative *ʔalik* 'send!' must be the imperative of an otherwise unattested verb **ʔalʔaka* 'to send',

[24]In several Quranic reading traditions these are still read *nabīʔ* and *nubūʔah*, as expected (Ibn Muǧāhid no date: 106–107).

[25]Read as *barīʔah* in several Quranic reading traditions (Ibn Muǧāhid no date: 693).

which has likewise irregularly lost its postconsonantal *ʔ*. Besides verbs, we may also see the irregular lack of representation of post-consonantal *ʔ* in other nouns, e.g. *malak* 'angel', which, considering its plural *malāʔikah* and etymological origin, was presumably originally **malʔak*.

The pseudo-verbs *niʕma* 'what a wonderful …' and *biʔsa* 'what an evil …', are presumably originally from **naʕima* and **baʔisa*, with vowel harmony and syncope. These original forms have disappeared from the classical language in their pseudo-verbal use, only retaining their verbal meaning: *naʕima* 'to be happy, glad' and *baʔisa* 'to be miserable, wretched'. However, other pseudo-verbs retain both unharmonized and unsyncopated forms as optional variants even in their pseudo-verbal use: *ḥasuna, ḥusna, ḥasna* 'how beautiful, magnificent', and *ʕaḏuma, ʕuḏma, ʕaḏma* 'how powerful, mighty'. Such syncopated and harmonized forms are claimed by the Arab grammarians themselves to be part of the eastern dialects, and absent in the Ḥiǧāzī dialects (Rabin 1951: 97), but surprisingly are retained for such pseudo-verbs.

Syncopated forms, while reported for regular verbs as well by the Arab grammarians (e.g. *šihda* or *šahda* for *šahida*), never show up in the Classical language. For some CaCiC nouns, syncopated forms are reported by lexicographers (e.g. *katf* and *kitf* besides *katif*), but it is not clear whether these syncopated forms are used in CA outside of these lexicons.

These kinds of dialectal forms that appear to have been incorporated into CA are indicative of the artificial amalgam that makes up the language, and require a much more in-depth discussion than the present chapter allows. It seems clear that the vast amount of dialectal variation that is described by the Arab grammarians, judiciously collected by Rabin (1951), does not end up in CA, but some amount of variants are either allowed, or are the only possible form present in the standard. The exact parameters that determine how and why such dialectal forms were incorporated into the language are currently unclear.

4 Conclusion

Due to CA and MSA being almost exclusively High literary registers, with no true native speakers, the type of language contact that we see in the Islamic period is rather different from what we may see in more natural language contact situations. We mostly see imposition of certain dialectal forms onto the Classical ideal. An interesting exception to this is the calquing of MSA words and phraseology upon "Standard Average European", where the speakers are dominant in neither the recipient nor the source language.

Borrowing can be detected in phonology, morphology and vocabulary from Greek, Aramaic and Ethio-Semitic from the Pre-Islamic period, which were then inherited by CA. In the Islamic period, it is most vocabulary that is borrowed, with a significant number of loans coming from Greek, Persian and Ottoman Turkish into CA.

Examining these Pre-Islamic borrowings, it has become clear that the Aramaic that has primarily influenced CA, contrary to what is popularly believed, was not a form of Syriac, but rather a more archaic variety. The historical implications of this have not yet been well-integrated into our understanding of pre-Islamic linguistic diversity in Arabia and neighbouring regions.

While some studies have looked at syntactic imposition of the spoken dialects onto MSA with promising results, this has not yet been applied to medieval texts written in CA. Nevertheless, considering the clear ethnic and geographic diversity of writers of CA, it seems likely that future work should be able to detect such influences even in the medieval period.

Further reading

Jeffery (2007) [1938] is still one of the most comprehensive books on loanwords in Quranic Arabic.
Hebbo (1984) is an in-depth study of foreign words as they appear in the Sīrah of Ibn Hišām.
Fraenkel (1886) is an in-depth discussion of Aramaic loanwords in Arabic, but in some respects outdated.
Nöldeke (1910) contains an important section on loanwords both from Arabic to the Ethio-Semitic languages and the other way around.
Blau (1969) is a pioneering work researching the interaction between European literary languages and the effects they have on the literary style of Modern Standard Arabic and Modern Hebrew.

The chapters on language contact in the *Encyclopaedia of Arabic Language and Linguistics* are also highly useful and informative, and contain many up to date references for contact with Greek (Gutas 2011), Persian (Asbaghi 2011), Aramaic Retsö (2011), and Turkish (Procházka 2011).

Acknowledgements

I thank Stefan Procházka, Christopher Lucas, Maarten Kossmann, Ahmad Al-Jallad for providing me with important references, comments and suggestions.

Abbreviations

*	reconstructed form
**	unattested form
1, 2, 3	1st, 2nd, 3rd person
ACC	accusative
Aram.	Aramaic
CA	Classical Arabic
CE	Common Era
DAT	dative
F	feminine
Gk.	Greek
Gz.	Gəʕəz
IMPF	imperfect (prefix conjugation)
Lat.	Latin
M	masculine
MSA	Modern Standard Arabic
NOM	nominative
OBL	oblique
PL	plural
PRF	perfect (suffix conjugation)
REL	relative pronoun
RL	recipient language
SG	singular
SL	source language
Syr.	Syriac
v.n.	verbal noun

References

Al-Jallad, Ahmad. 2015. *An outline of the grammar of Safaitic inscriptions*. Leiden: Brill.

Al-Jallad, Ahmad. 2017. Graeco-Arabica I: The southern Levant. In Ahmad Al-Jallad (ed.), *Arabic in context: Celebrating 400 years of Arabic at Leiden University*, 99–186. Leiden: Brill.

Al-Jallad, Ahmad. 2018. What is Ancient North Arabian? In Daniel Birnstiel (ed.), *Re-engaging comparative Semitic and Arabic studies*, 1–44. Wiesbaden: Harrassowitz.

Asbaghi, Asya. 2011. Persian loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Beeston, Alfred F. L. 1962. *A descriptive grammar of Epigraphic South Arabian*. London: Luzac.

Behnstedt, Peter. 2016. *Dialect atlas of North Yemen and adjacent areas*. Leiden: Brill. (Translated by Gwendolin Goldbloom).

Beyer, Klaus. 1984. *Die aramäischen Texte vom Toten Meer samt den Inschriften aus Palästina, dem Testament Levis aus der Kairoer Genisa, der Fastenrolle und den alten talmudischen Zitaten*. Vol. 1. Göttingen: Vandenhoeck & Ruprecht.

Blau, Joshua. 1969. *The renaissance of Modern Hebrew and Modern Standard Arabic: Parallels and differences in the revival of two Semitic languages*. Berkeley: University of California Press.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen. Band I: Laut- und Formenlehre*. Berlin: Reuther & Reichard.

Dichy, Joseph. 2011. Terminology. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Diem, Werner. 1980. Untersuchungen zur frühen Geschichte der arabischen Orthographie, II: Die Schreibung der Konsonanten. *Orientalia* 49. 67–106.

Dillmann, August. 2005. *Ethiopic grammar*. 2nd edn, revised by Carl Bezold and translated by James A. Crichton. Eugene, Oregon: Wipf & Stock. Originally published in 1907 by Williams & Norgate.

Eilers, Wilhelm. 1962. Iranisches Lehngut im arabischen Lexikon: Über einige Berufsnamen und Titel. *Indo-Iranian Journal* 5(3). 203–232.

Eilers, Wilhelm. 1971. Iranisches Lehngut im Arabischen. In *Actas, IV Congresso de Estudos Árabes e Islâmicos, Coimbra-Lisboa 1 a 8 de Septembro de 1968*, 581–657. Leiden: Brill.

Elmaz, Orhan. 2014. Investigating South Arabian words in al-Khalīl's Kitāb al-ʿAyn. *Supplement to the proceedings of the Seminar of Arabian Studies* 44. 29–42.

Elmaz, Orhan. 2016. *ʿarim* – A Sabaic word in the Qurʾān? In Roswitha G. Stiegner (ed.), *Süd-Arabien / South Arabia: A great "lost corridor" of mankind*, vol. 1, 213–244. Münster: Ugarit-Verlag.

Ferguson, Charles A. 1959. Diglossia. *Word* 15. 325–40.

Ferguson, Charles A. 1996. Epilogue: Diglossia revisited. In Alaa Elgibali (ed.), *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi*, 49–67. Cairo: American University in Cairo Press.

Fischer, Wolfdietrich. 2002. *A grammar of Classical Arabic*. 3rd, revised edn. Translated by Jonathan Rogers. New Haven: Yale University Press. Originally published in 1972 by Harrassowitz.

Fraenkel, Siegmund. 1886. *Die aramäischen Fremdwörter im Arabischen*. Leiden: Brill.

Glaß, Dagmar. 2011. Creating a modern standard language from medieval tradition: The Nahḍa and the Arabic academies. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 835–44. Berlin: De Gruyter Mouton.

Gutas, Dimitri. 1998. *Greek thought, Arabic culture: The Graeco-Arabic translation movement in Baghdād and early ʿAbbāsid society*. London: Routledge.

Gutas, Dimitri. 2011. Greek loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Gzella, Holger. 2015. *A cultural history of Aramaic: From the beginnings to the advent of Islam*. Leiden: Brill.

Hayajneh, Hani. 2011. The usage of Ancient South Arabian and other Arabian languages as an etymological source for Qurʾānic vocabulary. In Gabriel Said Reynolds (ed.), *New perspectives on the Qurʾān: The Qurʾān in its historical context 2*, 117–146. London: Routledge.

Hebbo, Ahmed. 1984. *Die Fremdwörter in der arabischen Prophetenbiographie des Ibn Hischām (gest. 218/834)*. Frankfut am Main: Peter Lang.

Huehnergard, John. 1997. *A grammar of Akkadian*. Atlanta: Scholars Press.

Huehnergard, John. 2005. Features of Central Semitic. In Agustinus Gianto (ed.), *Biblical and Oriental essays in memory of William L. Moran*, 155–203. Rome: Pontificio Istituto Biblico.

Huehnergard, John & Aaron Rubin. 2011. Modern Standard Arabic. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 259–278. Berlin: De Gruyter Mouton.

Ibn Muǧāhid, Abū Bakr. no date. *Kitāb al-sabʕah fī al-qirāʔāt*. Šawqī Ḍayf (ed.). Cairo: Dār al-Maʕārif.

Ibrahim, Zeinab. 2009. *Beyond lexical variation in Modern Standard Arabic: Egypt, Lebanon and Morocco*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Jacquart, Danielle (ed.). 1994. *La formation du vocabulaire scientifique et intellectuel dans le monde arabe*. Turnhout: Brepols Publishers.

Jeffery, Arthur. 2007. *The foreign vocabulary of the Qurʾān*. Leiden: Brill. Originally published in 1938 by Benoytosh Bhattacharyya.

Kossmann, Maarten. 2010. Parallel system borrowing: Parallel morphological systems due to the borrowing of paradigms. *Diachronica* 27(3). 459–487.

Kropftisch, Lorenz. 1977. Der französische Einfluß auf die arabische Schriftsprache im Maghrib. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 128. 39–64.

Leslau, Wolf. 1990. *Arabic loanwords in Ethiopian Semitic*. Wiesbaden: Harrassowitz.

Lucas, Christopher. 2015. Contact-induced language change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.

Mingana, Alphonse. 1927. Syriac influence on the style of the ḳurʾān. *John Rylands Library Bulletin* 11. 77–98.

Nöldeke, Theodor. 1910. *Neue Beiträge zur semitische Sprachwissenschaft*. Strasbourg: Karl J. Trübner.

Parkinson, Dilworth B. 2003. Future variability: A corpus study of Arabic future particles. In Dilworth B. Parkinson & Samira Farwaneh (eds.), *Perspectives on Arabic linguistics XV: Papers from the fifteenth Annual Symposium on Arabic linguistics, Salt Lake City 2001*, 191–211. Amsterdam: John Benjamins.

Parkinson, Dilworth B. 2007. Sentence subject agreement variation in Arabic. In Zeinab Ibrahim & Sanaa A. M. Makhlouf (eds.), *Linguistics in an age of globalization: Perspectives on Arabic language and teaching*, 67–90. Cairo: American University in Cairo Press.

Parkinson, Dilworth B. & Zeinab Ibrahim. 1999. Testing lexical differences in regional standard Arabic. In Elabbas Benmamoun (ed.), *Perspectives on Arabic linguistics XII: Papers from the twelfth Annual Symposium on Arabic linguistics*, 183–202. Amsterdam: John Benjamins.

Procházka, Stephan. 2011. Turkish loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

van Putten, Marijn. 2018. Hamzah in the Quranic consonantal text. *Orientalia* 87(1). 93–120.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Ratcliffe, Robert R. 1998. *The broken plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology*. Amsterdam: John Benjamins.

Retsö, Jan. 2011. Aramaic/Syriac loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Rubin, Aaron D. 2010. *The Mehri language of Oman*. Leiden: Brill.

Ryding, Karin C. 2011. Modern Standard Arabic. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 844–50. Berlin: De Gruyter Mouton.

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.

Stein, Peter. 2011. Ancient South Arabian. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1042–1073. Berlin: De Gruyter 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.

Weninger, Stefan. 2009. Der Jemen als lexikalisches Ausstrahlungszentrum in der Antike. In Werner Arnold, Michael Jursa, Walter W. Müller & Stephan Procházka (eds.), *Philologisches und Historisches zwischen Anatolien und Sokotra: Analecta semitica in memoriam Alexander Sima*, 395–410. Wiesbaden: Harrassowitz.

Weninger, Stefan. 2011a. Ethio-Semitic in general. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1114–1123. Berlin: De Gruyter Mouton.

Weninger, Stefan. 2011b. Old Ethiopic. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1124–1142. Berlin: De Gruyter Mouton.

Whorf, Benjamin Lee. 1956. *Language, thought and reality: Selected writings of Benjamin Lee Whorf*. John B. Carroll (ed.). Cambridge, Massachusetts: Technology Press of MIT.

Wilmsen, David. 2010. Dialects of written Arabic: Syntactic differences in the treatment of object pronouns in Egyptian and Levantine newspapers. *Arabica* 57. 99–128.

Chapter 3

Arabic in Iraq, Syria, and southern Turkey

Stephan Procházka
University of Vienna

This chapter covers the Arabic dialects spoken in the region stretching from the Turkish province of Mersin in the west to Iraq in the east, including Lebanon and Syria. The area is characterized by a high degree of linguistic diversity, and for about two and a half millennia Arabic has come into contact with various other Semitic languages, as well as with Indo-European languages and Turkish. Bilingualism, particularly with Aramaic, Kurdish, and Turkish, has resulted in numerous contact-induced changes in all realms of grammar, including morphology and syntax.

1 Current state and historical development

The region discussed in this chapter is linguistically extremely heterogeneous: in it three different Arabic dialect-groups, plus several other languages, are spoken. The two main Arabic dialect-groups are Syrian and Iraqi, the distribution of which does not exactly correspond to the political boundaries of those two countries. Syrian-type dialects are also spoken in Lebanon, in three provinces of southern Turkey (Mersin, Adana,[1] Hatay), and in one village on Cyprus (cf. Walter, this volume). In Iraq, Arabic is mainly spoken in Mesopotamia proper, whereas considerable parts of the mountainous parts of the country are Kurdish-speaking. Arabic dialects which are very akin to the Iraqi ones extend into northeastern Syria and southeastern Anatolia (for the latter see Akkuş, this volume). These two groups are geographically divided by a third dialect-group, which came here

[1]The dialects spoken in Mersin and Adana provinces will henceforth referred to as Cilician Arabic.

Stephan Procházka. 2020. Arabic in Iraq, Syria, and southern Turkey. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 51–82. Berlin: Language Science Press. DOI:??

with an originally (semi-) nomadic population from northern Arabia. Today, this variety preponderates in all villages and most towns between the eastern outskirts of Aleppo and the right bank of the Tigris, and stretching north into the Turkish province of Şanlıurfa.

The total number of native Arabic speakers in the whole region is estimated to be 54 million (see Table 1). The dialects of large urban centers like Beirut, Damascus, Aleppo, Baghdad, and Mosul have become supra-regional prestige varieties that are also used in the media and therefore understood by most inhabitants of the respective countries. The situation is very different in Turkey, where the local Arabic is in sharp decline and public life is exclusively dominated by Turkish. Only recently has the position of Arabic in Turkey been socially enhanced by the influx of more than 3.5 million Syrian refugees fleeing the civil war that started in 2011.[2]

Table 1: Speaker populations for dialects of Arabic

Country	Speakers
Syria	17,000,000
Lebanon	6,000,000
Iraq	30,000,000
Turkey	1,000,000

Arabic was spoken in the region long before the advent of Islam (Donner 1981: 95) but became the socially dominant language in the wake of the Muslim conquests in the seventh century CE. From that time until the end of the tenth century, when Bedouin tribes seized large parts of central and northern Syria, there was probably a continuum of sedentary-type dialects that stretched from Mesopotamia to the northeastern Mediterranean (Procházka 2018: 291). During the Mongol sacking of Iraq in 1258, much of the population was killed or expelled. This resulted in far-reaching demographic and linguistic changes as the original sedentary-type dialects could merely hold ground in Baghdad and the larger settlements to its north. Further south they persisted only among the non-Muslim population. Most of today's Iraq was re-populated by people who spoke Bedouin-type dialects (mostly coming from the Arabian Peninsula), which over the centuries have heavily influenced the speech of even most large cities (Holes 2007). Very similar dialects are spoken further south and in the Iranian province

[2]See UNHCR figures at https://data2.unhcr.org/en/situations/syria/location/113.

of Khuzestan (see Leitner, this volume). The foundation of nation states after World War One caused a significant decrease in contact between the different dialect groups and an almost complete isolation of the Arabic dialects spoken in Turkey.

2 Contact languages

During its two-and-a-half-millennia presence in the region, Arabic has come into contact with many languages, both Semitic and non-Semitic. Those most relevant for the topic will be treated in more detail below (for Syria, see also Barbot 1961: 175–177). Akkadian was spoken in southern Iraq until about the turn of the eras, i.e. the first century CE.[3] Greek was the language of administration in Greater Syria until the Arab conquest (Magidow 2013: 185–187) and continued to play a role for Orthodox Christians.[4] During Crusader times, Arabic speakers in Syria came into contact with various medieval European languages; and along the Mediterranean coast the so-called Lingua Franca (see Nolan, this volume) was an important source for the spread of particularly nautical vocabulary for many centuries (Kahane et al. 1958). Since the nineteenth century, locally restricted contacts between Arabic and Armenian and Circassian have existed in parts of Syria and Lebanon.

2.1 Aramaic

Aramaic is a Northwest Semitic language and thus structurally very similar to Arabic. Different varieties of Aramaic were the main language in Syria and Iraq from the middle of the first millennium BCE and it can be assumed that some contact with Arabic existed even at that time. From the first century CE onwards, the southern fringes of the Fertile Crescent became largely Arabic-dominant and there was significant bilingualism with Aramaic, particularly in the towns along the edge of the steppe, such as Petra, Palmyra, Hatra, and al-Ḥīra (Procházka 2018: 260–262). Though after the Muslim conquests Arabic eventually became the majority language, it did not oust Aramaic very quickly: the historical sources suggest that Aramaic dominated in the larger towns and the mountainous regions of Syria and Lebanon for a long time. In Iraq, by contrast, the massive influx of Arabs into the cities fostered their rapid Arabization, while Aramaic continued

[3]For Akkadian lexical influence on Arabic, see Holes (2002) and Krebernik (2008).

[4]The enormous influence of Modern Greek on the Arabic spoken in the Kormakiti village of Cyprus is discussed by Walter (this volume). For a detailed study, see also Borg (1985).

to be spoken in the countryside (Magidow 2013: 184; 188). But over the centuries, the diverse Aramaic dialects became marginalized and, with very few exceptions, were finally relegated to non-Muslim religious minorities, particularly Christians and Jews, in peripheral regions like Mount Lebanon and the Anti-Lebanon Mountains, where Aramaic was prevalent until the eighteenth century (Retsö 2011). Western Aramaic is still spoken in three Syrian villages, the best known of which is Maaloula.[5] There also remain speakers of Neo-Aramaic in northern Iraq.[6]

It is hard to establish the degree of bilingualism in the past, but it can be assumed that it was mostly Aramaic L1 speakers who had a command of Arabic and not vice versa. In the present time, nearly all remaining Aramaic speakers in Syria are fluent in Arabic. In Iraq this is mainly true of those living in the plain just north of Mosul (Arnold & Behnstedt 1993; Coghill 2012: 86). The influence of different strata of Aramaic on spoken Arabic is a long debated issue, various scholars rating it from considerable to negligible (Hopkins 1995: 39; Lentin 2018).

2.2 Persian and Kurdish

For many centuries, Arabic and the two Western Iranian languages Persian and Kurdish have influenced each other on different levels. Persian-speaking communities existed in medieval Iraq, and economic and cultural contacts between Mesopotamia and Iran have continued to the present (cf. Gazsi 2011). An important factor of language contact are the holy shrines of the Imams in Kerbela, Najaf, and other Iraqi cities, which have always attracted tens of thousands of Persian speaking Shiites every year. Intensive contacts between speakers of Kurdish and Arabic have existed since at least the tenth century, particularly in Northern Iraq, northeast Syria, and southeast Anatolia (see Akkuş, this volume). Until their exodus in the early 1950s, the Arabic-speaking Jewish communities which existed in Iraqi Kurdistan usually had a native-like command of Kurdish (Jastrow 1990: 12). Due to the multilingual character of the region, bilingualism in Kurdish and Arabic is still relatively widespread, particularly in urban settings, though with Kurds usually much more fluent in Arabic than the other way around.[7] However, for obvious reasons, little linguistic research has been done in Iraq for decades, which makes it impossible to give up-to-date information about the linguistic situation in ethnically-mixed cities like Kirkuk.

[5]The village heavily suffered from the jihadist occupation of 2013–2014, but after government troops had retaken control over the region, many inhabitants returned and began its reconstruction (cf. the reports collected at http://friendsofmaaloula.de/).

[6]See Coghill (2012) and http://glottolog.org/resource/languoid/id/nort3241.

[7]With significant exceptions in some parts of southeast Anatolia; see Akkuş (this volume).

2.3 Ottoman and Modern Turkish

Contacts between spoken Arabic varieties and various Turkic languages existed from the ninth century onwards. These early contacts, however, left hardly any traces in Arabic except for a handful of loanwords. In the sixteenth century, the Ottomans established their rule over most Arab lands, including Syria, Lebanon, and Iraq. This domination lasted four hundred years, until World War One. Particularly in the provinces of Aleppo and Mosul, there were a relatively high percentage of Turkish speakers and probably a significant degree of bilingualism.[8] As the language of the ruling elite, Turkish had high prestige and therefore was at least rudimentarily spoken by many inhabitants of those regions, especially urban men. The collapse of the Ottoman Empire put an abrupt end to Turkish–Arabic contacts, which today remain intensive only among the Arabic varieties spoken within the borders of Turkey itself, where most Arabic speakers are fluent in Turkish, the dominant language in all contact settings.

In some areas of Syria and in northern Iraq, the Arabic-speaking population lives side by side with several hundred thousand speakers of Turkish and Azeri Turkish, who call themselves Turkmens. Unfortunately, no reliable data on the sociolinguistic settings and the degree of bilingualism exist for those areas. Again, it can be assumed that most of the Turkmens in both countries are dominant in Turkish, but know Arabic as a second language.

2.4 French and English

After World War One, Syria and Lebanon stayed under the French mandate and Iraq under the British mandate until they reached independence.[9] French is still widely spoken as a second language in Lebanon, especially by Christians. In Iraq, English has maintained its position as by far the most important foreign language – a fact which was reinforced by the US military occupation from 2003 to 2010.

2.5 Intra-Arabic contacts

Contacts between different Arabic varieties, for instance between speakers of rural and urban dialects, happen on an everyday basis and often trigger short-term accommodation without leading to long-lasting changes. The situation is different with regard to the enduring contacts between the Bedouin and the sedentary

[8] See Wilkins (2010: xv) for Aleppo. Koury (1987: 103) maintains that Aleppo's hinterland was culturally even more Turkish than Arab. For Mosul, see Shields (2004: 54–55).

[9] Iraq in 1932, Lebanon in 1943, Syria in 1946.

populations, whose dialects considerably differ from each other.[10] Such contacts are most intense at the periphery of the Syrian steppe and along the middle Euphrates, where scattered towns with sedentary dialects like Palmyra, Deir ez-Zor and Hit are surrounded by an originally nomadic population. Though the nomadic way of life has been abandoned by most of them, they still speak Bedouin-type Arabic dialects. As the nomads were, for many centuries, both socially and economically dominant, speakers of sedentary dialects often adopted linguistic features from the more prestigious Bedouin (though reverse instances are also found; cf. Behnstedt 1994b: 421). Due to the historical circumstances mentioned in §1, Bedouins also had a strong linguistic impact on Iraqi dialects. In Baghdad the sedentary dialect of the Muslim population has been gradually Bedouinized due to massive migration from the countryside to the city (Palva 2009). The Christian and, in former times, Jewish inhabitants, on the other hand, preserved their original sedentary-type dialects because they had much less contact with the Muslim newcomers.

3 Contact-induced changes

Change induced by contact with Aramaic almost exclusively happened through imposition, that is, by Aramaic speakers who had learned Arabic as a second language and later often completely shifted to Arabic. This explains the relatively numerous phonological changes and pattern replications in syntax. Lexical transfers from Aramaic certainly were also made by Arabic-dominant speakers, particularly in semantic fields like agriculture that included novel concepts for the mostly animal-breeding Arabs.

The same is true for transfers from Greek, for which a very low level of bilingualism can be assumed. Thus we find only matter replication in the form of loanwords, mostly in domains where lexical gaps in older layers of spoken Arabic are likely.

In the case of Kurdish, bilingualism is much more widespread among speakers of the source language, suggesting imposition. This might explain the phonological changes, as speakers dominant in the source language tend to preserve its phonological features (Lucas 2015: 532). The relatively small number of instances of lexical matter replication is probably the result of the fact that Arabic has long been regarded as the more prestigious by speakers of both the source and the

[10] Since these two speech communities differ from each other in so many ways, it is a relatively robust approach to rate the following features as results of dialect contact and not mere variation (cf. Lucas 2015: 533).

recipient language.

The numerous loanwords from Persian into Iraqi Arabic may well be the result of matter replication by agents who were dominant in the recipient language Arabic. Starting with the rule of the Abbasid caliphs in the eighth century CE and continuing to the present, Iranian material culture and cuisine often had a great impact on neighboring Mesopotamia. There were also many intellectuals, among them praised writers of Arabic prose, who were actually Iranians and hence knew both languages. Frequent contacts on the everyday level caused additional borrowing of ordinary vocabulary and the retention of sounds that are replaced in Persian loans found in Classical Arabic or other dialects.[11]

Changes induced by contact with Ottoman Turkish may have happened mostly through Arabic-dominant speakers. The current situation of Arabic speakers in Turkey is, however, very different, because at least the last two generations have acquired Turkish as an L2 or even as a second L1 at very young age. Thus, at least some of the contact phenomena described in the following paragraphs may be examples of linguistic convergence (see Lucas 2015: 525).

French and English have largely remained typical "foreign languages" learned at school or in business with a considerable amount of bilingualism only in some urban settings of Lebanon, particularly Beirut. The agents of change are certainly dominant in the recipient language.

The distinction between the two transfer types is not always clearly discernible in case of intra-Arabic contact-induced changes. In the towns of the Syrian steppe and the middle Euphrates the agents of change were mostly the sedentary population who adapted their speech towards the norms of the socially more prestigious Bedouin. However, there has always been inter-marriage, and Bedouins often settled in towns and may well have adopted features from the local sedentary variety. Especially in cases like Muslim Baghdadi (see §1) we may assume with good reason that the Bedouin character of today's variety developed through both imposition and borrowing.

3.1 Phonology

3.1.1 Aramaic-induced changes

It has been hypothesized that several phonological features of the Syrian and Lebanese dialects are due to contact-induced influence of Aramaic. But in the case of the shift from interdental fricatives to postdental plosives (/ð/ > /d/; /θ/

[11]The phonological changes are not, however, only the result of Persian influence (cf. §2.1).

> /t/; /ð/ > /ḍ/) this is unlikely because: (i) this sound change is common cross-linguistically; (ii) it does not occur in all dialects of the region; and (iii) it is found in many other Arabic dialects without an Aramaic substrate.

A phonotactic characteristic of most dialects spoken along the Mediterranean, from Cilicia in the north to Beirut in the south, is that all unstressed short vowels (including /a/) in open syllables are elided,[12] whereas in other dialects east of Libya only /i/ and /u/ in this position are consistently dropped.

(1) Cilician Arabic (Procházka 2002b: 31–32; 130)
Old Arabic (OA) *raṣāṣ > *ṛṣāṣ* 'lead, plumb'
OA *miknasa > *mikinsi* 'broom'[13]
*fataḥ-t > *ftaḥt* 'I opened'

Because this rule corresponds to the phonotactics of Aramaic and is otherwise not found in the same degree except in Maghrebi dialects (cf. Benkato, this volume), pattern replication is likely, though cannot be proved.[14]

In roughly the same region, except Cilicia and many dialects of Hatay,[15] the diphthongs /ay/ and /aw/ are only preserved in open syllables, but monophthongized to /ē/ and /ō/ respectively in closed syllables. In some regions, for instance on the island of Arwad, both diphthongs merge to /ā/ in closed syllables (Behnstedt 1997: map 31).

(2) Arwad, western Syria (Procházka 2013: 278)
OA *bayt, *baytayn > *bāt, baytān* 'house, two houses'
OA *yawm, *yawmayn > *yām, yawmān* 'day, two days'
OA *bayn al-iθnayn > *bān it-tnān* 'between the two'

Likewise, in older layers of Aramaic, diphthongs were usually monophthongized in closed syllables (for Syriac see Nöldeke 1904: 34), which makes imposition by L1 speakers of Aramaic rather likely (Fleisch 1974b: 227).

Another striking phenomenon is the split of historical /ā/ into /ō/ and /ē/ that is found in scattered areas of the Levant, particularly northern Lebanon, around the Syrian port of Tartous, the Qalamūn Mountains, and the exclusively Christian

[12] Therefore, Cantineau (1960: 108) called them *parlers non différentiels* – a term still very often applied in Arabic dialectology – as they make no distinction in the treatment of the three short vowels.

[13] With insertion of an epenthetic /i/ to avoid a sequence of three consonants.

[14] Cf. Diem (1979: 47); Arnold & Behnstedt (1993: 69–71); Weninger (2011: 748).

[15] Where this phenomenon occurs only in Alawi villages (Arnold 1998: 84).

town of Maḥarde on the Orontes River.[16] Because in many varieties of Aramaic the old Semitic /ā/ is reflected as /ō/, it could be assumed that Aramaic speakers transferred their peculiar pronunciation to Arabic when learning it. Fleisch (1974a: 49) rejected the hypothesis of an Aramaic influence, arguing that the conditioned distribution of the two allophones is merely a further development of the [ɒ] : [æ] split widely attested for Lebanon and parts of western Syria. However, in the Syrian Qalamūn Mountains there are dialects with an unconditioned shift (Behnstedt 1992), and this is precisely the region where the shift from Aramaic to Arabic occurred relatively late, probably after a long phase of bilingualism. In the town of Nabk, for instance, one can infer that the former Aramaic speaking inhabitants would have simply turned every /ā/ into /ō/ – except those which long before had become [ē] (or [ɛ̄]) as a result of the so-called conditioned *imāla* (i.e. the tendency of long /ā/ to be raised towards [ē] or even [ī] if the word contains an /i/ or /ī/).[17] Example (3) clearly shows that the distribution of the allophones is not conditioned by the consonantal environment.

(3) Nabk, Syria (Gralla 2006: 20)
OA *ṭābiḥ > *ṭēbeḥ* 'cooking' vs. OA *ṭālib > *ṭōleb* 'student'
OA *ḥāmil > *ḥēmel* 'pregnant' vs. OA *ḥāmiḍ > *ḥōmeḍ'* 'sour'

In these cases Aramaic influence seems plausible. For the region of Tripoli it may be assumed that Aramaic bilinguals from the adjacent mountains used [ō] instead of [ā] when speaking Arabic and thus reinforced the already existing [ɒ] : [æ] split.[18]

3.1.2 The "new" phonemes /č/, /g/, and /p/

Consonantal phonemes that are originally alien to Arabic are found in all Arabic dialects spoken in Turkey (see also Akkuş, this volume), northern Syria, and Iraq. These are the unvoiced affricate /č/, the voiced /g/,[19] and the unvoiced /p/, the latter mainly used in Iraq. The emergence of these sounds was very likely

[16] For details cf. Behnstedt (1997: map 32). The conditioned shift /ā/ > /ō/ is also found in and around Tarsus in Turkey (Procházka 2002b: 37–38).

[17] Cf. Arnold & Behnstedt (1993: 68).

[18] For discussion see Fleisch (1974a: 48–50; 1974b: 133–136), Diem (1979: 45–46); Behnstedt (1992); Arnold & Behnstedt (1993: 67–68); Weninger (2011: 748).

[19] The sound [g] is prevalent in whole Syria and Lebanon but seems to have phonemic status only in the north (Sabuni 1980: 26). For further examples and discussion see Ferguson (1969). This "foreign" /g/ must therefore be differentiated from the /g/ which is the regular reflex of OA *q. The latter development is found in many Bedouin-type dialects.

contact-induced, but it is often impossible to discern which language triggered each development: all three sounds are found in Persian, Kurdish, Turkish, and the Lingua Franca. For the dialects of Cilicia, Hatay and Syria, the main source language doubtless was Turkish. The sound /p/ in the Iraqi dialects was probably first introduced through contact with Persian and Kurdish, and then reinforced by Ottoman Turkish. In the Bedouin-type dialects of the region, the phonemes /č/ and /g/ are not products of contact-induced change but occur due to internal sound changes, unvoiced /č/ as a conditioned affricated variant of /k/ and /g/ as the ordinary reflex of OA *q.

Thus, it can be assumed that over the centuries speakers of the sedentary dialects of Iraq and Syria borrowed either from other languages or from Bedouin Arabic varieties words that possess these two sounds, which subsequently were fully incorporated into the phonemic inventory. This development may have been facilitated by the fact that the three sounds /č/, /p/, and /g/ are not fundamentally unfamiliar to Arabic, but are the voiceless/voiced counterparts of the well-established phonemes /ǧ/, /b/, and /k/. It seems no accident that the new sound /č/ is much more often found in dialects that have preserved the affricate /ǧ/ than in those where it has shifted to /ž/, as illustrated in examples (4) and (5).

(4) Aleppo (Sabuni 1980: 205–210)
čanṭāye 'handbag' (Turkish *çanta*)
čwāl 'sack' (Turkish *çuval*)
čāy 'tea' (Turkish *çay*)
gağaleg 'nightgown' (Turkish *gecelik*)

The words given in (4) are usually pronounced with [š] instead of [č] in the central Syrian and Lebanese dialects where contact with Turkish was less intense and /ǧ/ is reflected as /ž/.[20]

(5) Mosul (own data)
ṣūč 'fault' (Turkish *suç*)
pāča 'stew of sheep and cow legs and innards' (Kurdish/Persian *pāče*)
zangīn 'rich' (Turkish *zengin*)

Once integrated into the phonological system, these sounds not only enabled easier integration of loanwords from other languages like French and English

[20] Cf. Behnstedt (1997: maps 18, 19, 25). For details and more examples see Sabuni (1980: 205–210), who lists all words with *č/g* in Aleppo, and Procházka (2002a: 185) for Cilician Arabic.

(see §2.4), but sometimes also resulted in the spread of assimilation-induced allophones from single words to the whole paradigm or even root. In Aleppo one finds *yəkdeb > *yəgdeb* 'he lies' due to assimilation. The *g* subsequently was transferred to other words derived from the root: *gadab* 'he lied', *gədbe* 'lie', and *gaddāb* 'liar' (Sabuni 1980: 26, 209).

Speakers of sedentary dialects who had everyday contact with Bedouins – for example the inhabitants of Deir ez-Zor and Khatuniyya – first integrated /č/ and /g/ into their phonemic inventory through the borrowing of typically Bedouin vocabulary such as *dabča* 'a Bedouin dance' (Khawetna; Talay 1999: 29) and *ṭabga* 'milk-bowl' (Soukhne; Behnstedt 1994a: 310). These sounds then entered other fields of the lexicon, which led to unpredictable distribution, including doublets, as in (6)–(8).

(6) Khawetna (Talay 1999: 28–31)
gəṣṣa 'forehead', but *qəṣṣa* 'story' (OA *quṣṣa / *qiṣṣa)
dīč 'rooster' (OA *dīk)

(7) Deir iz-Zor (Jastrow 1978: 42–43).
gāʕ 'soil' (OA *qāʕ)
čam 'how much?' (OA *kam)

(8) Baghdad (Palva 2009: 18–19)
guffa 'large basket' (OA *quffa), but *quful* 'lock' (OA *qufl)
ʕigab 'to pass', but *ʕiqab* 'to follow' (both OA *ʕaqab)

The opposition /k/ : /č/ has even entered morphology, particularly with the 2SG suffixes: *ʔabū-k* 'your (SG.M) father' vs. *ʔabū-č* 'your (SG.F) father'. In the Syrian oasis of Soukhne, long-term contact with speakers of Bedouin dialects caused a chain of phonetic changes: first /k/ shifted to /č/, which originally was the reflex of OA /ǧ/; then /č/ (< /ǧ/) shifted further to /ts/, which has become a unique feature of the local dialect. The unconditioned shift from /k/ > /č/, which is not found in the Bedouin dialects, in turn caused a shift from /q/ > /k/.[21]

(9) Soukhne (Behnstedt 1994a: 226, 344, 357, 360)
kirbi 'water-skin' (< OA *qirba, Bedouin *girba*)
čalb 'dog' (< OA *kalb, Bedouin *čalib*)
čurr 'donkey foal' (< OA *kurr, Bedouin *kuṛṛ*)
tsubn 'cheese' (< OA *ǧubn, Bedouin *ǧubun*)

[21]See Behnstedt (1994a: 4–11) for details.

3.2 Morphology

3.2.1 Diminutive

The Aramaic diminutive suffix *-ūn* has become restrictedly productive in Iraqi Arabic (Masliyah 1997: 72), as illustrated in (10). In Syria and Lebanon it is only found in fossilized forms such as *šalfūn* 'young cockerel' and *qafṣūne* 'little cage'. Such kinds of morphological transfer are usually triggered by lexical borrowing. Thus, it may be assumed that this suffix spread from loanwords like *šalfūne* 'small knife blade' < Aramaic *šelpūnā* 'little knife' (cf. Féghali 1918: 82).[22]

(10) Iraq (Masliyah 1997: 72)
darb 'road' > *darbūna* 'alley'
gṣayyir 'short' < *gṣayyrūn* 'very short'
mḥammdūn hypocoristic form of the name *Muḥammad*

3.2.2 Morphological templates

Syrian and Lebanese dialects exhibit a few word patterns (templates) that are attested for OA (and other dialects) but seem to have become widespread through contact with Aramaic due to their frequency in the latter. These are the verbal pattern $\text{ša}C_1C_2\text{a}C_3$ and the (primarily diminutive) nominal patterns $C_1\text{a}C_2C_2\text{ū}C_3$ and $C_1\text{a}C_2C_3\text{ū}C_4$.[23]

An example of the first is *šanfaḥ* 'to puff up', related to *nafaḥ* 'to blow up' (Féghali 1918: 83; cf. Lentin 2018: 201 for further discussion); the nominal forms are illustrated in (11) and (12).

(11) Aleppo (Barthélemy 1935: 104, 158, 851)
ǧaḥḥūš 'little donkey' (related to *ǧaḥš* 'young donkey')
ḥassūn 'goldfinch' (related to the personal name *ḥasan*)
namnūme 'small louse' (cf. *naml* 'ants')

The pattern $C_1\text{a}C_2C_2\text{ū}C_3$(i) is still productive in the whole region, including the Bedouin dialects, to derive hypocoristic forms from personal names:

(12) *fāṭma* > *faṭṭūma*
ḥalīme > *ḥallūma*
aḥmad/*mḥammad* > *ḥammūdi*

[22] This must be a very old borrowing because the suffix is also found in the Gulf dialects (e.g. *ḥabbūna* 'a little' Holes 2002: 279) and even in Tunisian Arabic (Singer 1984: 496), where direct Aramaic influence can be excluded.

[23] For the latter two see Corriente (1969) and Procházka (2004).

3.2.3 Pronouns

In all Syrian and Lebanese dialects, as well as in Anatolia, the 2PL and 3PL pronouns exhibit an /n/ in place of the /m/ that is found in other Arabic dialects, which makes them look as if they were reflexes of OA feminine forms (Table 2).

Table 2: 2PL and 3PL pronouns

	Damascus	Jerusalem	OA PL.F	Syriac PL.M
2PL	*ʔəntu / -kon*	*ʔintu/-kom*	*ʔantunna / -kunna*	*ʔatton / -kon*
3PL	*hənne(n) / -hon*	*humme / -hom*	*hunna / -hunna*	*hennon / -hon*

Because generalization of the feminine is unlikely,[24] these forms have often been explained as a contact-induced change. In Aramaic the corresponding pronouns also have /n/ (for Syriac see Muraoka 2005: 18). In particular, the 3rd person forms with final *-n* exactly mirror the Aramaic pattern, but lack a plausible intra-Arabic etymology. Thus imposition seems plausible. Nevertheless, substratum influence has been doubted, particularly because of the infrequent evidence of *n*-pronouns in other regions.[25]

3.2.4 Vocative suffixes

The suffixes *-o* (in the west of the region) and *u* (in the east) can be attached to various kinship terms and given names when used for direct address, usually hypocoristically.[26]

(13) Urfa (own data)
šnōnak ḫayy-o? 'Brother, how are you?'
ǧidd-o 'Grandfather!'
ʕamm-o '(paternal) Uncle!'
ḫāl-o '(maternal) Uncle!'

[24]This is mainly because the feminine forms are only used for addressing groups of females, whereas the masculine forms may also refer to a mixed group. Therefore, the masculine forms are certainly more frequent. In all Arabic dialects except those mentioned above, the gender-neutral plural forms are clearly derived from the historical masculine.

[25]See Owens (2006: 244–245) and Procházka (2018: 283–284) for details.

[26]See also Ferguson (1997: 187).

In Syria the suffix is also added to female nouns: *ʕamm-t-o* '(paternal) Aunt!' and *ḫāl-t-o* '(maternal) Aunt!', whereas in Iraq the corresponding forms end in *-a*: *ʕamm-a, ḫāl-a.*

Since this suffix has no overt Arabic etymology, it has been assumed to be a borrowing of the Kurdish vocative *-o* (e.g. Grigore 2007: 203). The Persian suffix *-u* also forms affective diminutives,[27] which would make Persian influence possible, at least for Iraq.[28] However, the distribution of this feature extends far beyond even indirect contact with Kurdish or Persian,[29] though reinforcement and influence on the phonology may be possible for certain regions. Similar endings in Aramaic (Fassberg 2010: 88–89) and Ethiopian (Brockelmann 1928: 122) suggest a common Semitic origin (see also Pat-El 2017: 463–465).

3.2.5 Turkish derivational suffixes

All dialects of the region have incorporated the Turkish suffix *-ci* [ʤi] into their nominal morphology, as illustrated in (14) and (15). This suffix has become productive and is therefore a good example of morphological matter borrowing (Gardani et al. 2015). It is widely used for expressing professions, occupations, and habitual actions – the latter overwhelmingly pejorative, or at least humorous. In Iraqi dialects the suffix is reflected as *-či*, which corresponds to its pronunciation in the regional Turkish varieties. In the other varieties, it follows the usual development of *ğ, which means that it is realized as *-ği* or *-ži.*

(14) Syria/Damascus (own data)
kahrab-ži 'electrician' (*kahraba* 'electricity')
nəswān-ži 'womanizer' (*nəswān* 'women')
maškal-ži 'troublemaker' (*məšəkle* 'problem')

(15) Iraq (Procházka-Eisl 2018: 40–44)
pančar-či 'tire repairman' (*pančar* 'puncture')
mharrib-či 'human trafficker' (*mharrib* 'one who helps s.o. to escape')
ʕarag-či 'drunkard' (*ʕarag* 'aniseed brandy')

[27] E.g. *pesar-u* 'kid'; *ʕamm-u* is even the common word for 'uncle' (Perry 2007: 1011).

[28] In the Iraqi dialects the vowel is *-u*, e.g. *ʕamm-u, ḫāl-u* and *ğidd-u* (Abu-Haidar 1999: 145).

[29] The suffix is, for instance, attached to given names for endearment in the Gulf dialects, cf. Holes (2016: 128). The address forms *ya ʕamm-u, ya ḫāl-u* 'uncle', *gidd-u* 'grandfather', *sitt-u* 'grandmother' are used in Cairo, where hypocoristic variants of given names are likewise attested, e.g. *mīšu* for *hišām* (Woidich 2006: 109). The suffix *-o/-u* in address forms is also attested in eastern Sudan (Stefano Manfredi, personal communication), and in the Maghreb; Prunet & Idrissi (2014: 184) provide a list of such nouns for Morocco.

The suffix clearly fills a morphological gap, because it enables morphologically transparent derivation even from loanwords, by preserving the basic, immediately recognizable word – in contrast to the Arabic $C_1aC_2C_2āC_3$ pattern or participles, which are derived from the root (for details see Procházka-Eisl 2018).

To a lesser extent other Turkish suffixes have enhanced the morphological devices of the dialects treated here,[30] specifically the relative suffix *-li*, the privative suffix *-siz*, and the abstract suffix *-lik*, which is reflected as *-loɣiyya* in Iraq, i.e. with the Arabic abstract morpheme affixed. For the most part these suffixes appear in Turkish loanwords, e.g. Cilicia *ṣiḥḥat-li* (< Turkish *sıhhatlı*) 'healthy', *raḥaṭ-ṣīz* (< Turkish *rahatsız*) 'uncomfortable'. Only in Iraq have they gained a certain degree of productivity, particularly *-sizz* and *-loɣiyya*:

(16) Iraq (Masliyah 1996: 293–294)
muḥḥ-sizz 'stupid, brainless'
ḥaya-sizz 'shameless'
ḥaywān-loɣiyya 'ignorance' (lit. 'animal-ness')
zmāl-loɣiyya 'stupidity' (lit. 'donkey-ness')

3.2.6 Light-verb constructions

Arabic dialects spoken in Turkey not infrequently use light-verb constructions (in Turkish grammar mostly called phrasal verbs) which consist of the verb 'to do' plus a following noun (see also Akkuş, this volume). Such compound verbs are very frequent in Turkish (and Kurdish) and enable easy integration of foreign vocabulary into the verbal system. The light verbs found in the Arabic dialects show that this formation is a case of selected pattern replication because, first, not all examples are exact copies of the Turkish model, and second, the word order follows the Arabic VO rather than the Turkish OV pattern:

(17) Harran–Urfa (own data)
sāwa qaza (Turkish *kaza yapmak*) 'to have an accident'
sāwa ʕēš (in Turkish not a phrasal verb, but *pişirmek*) 'to cook'

(18) Cilician Arabic (Procházka 2002b: 198)
sawwa zarar (Turkish *zarar vermek*) 'to harm'
sawwa ḫayir (Turkish *hayır işlemek*) 'to do a good deed'

[30] See Halasi-Kun (1969: 68–71); Sabuni (1980: 168); Masliyah (1996); Procházka (2002b: 186).

3.2.7 Intra-Arabic dialect contact

Concerning intra-Arabic contact, here we see that this has led to the adoption of typical Bedouin-type pronouns into sedentary dialects (cf. Palva 2009: 27–29), e.g.:

(19) Baghdad, Deir ez-Zor, Soukhne
ʔəḥna for *nəḥna* 1PL

(20) Baghdad
ʔāni for *ʔana* 1SG

In addition, as shown in Table 3, virtually all the eastern sedentary dialects of Syria have copied the typical Bedouin-type active participles of the verbs 'to eat' and 'to take', which exhibit initial *m-* (Behnstedt 1997: map 175).

Table 3: Active participles of the verbs 'to eat' / 'to take'

Bedouin	Soukhne	Palmyra	Damascus
māčil / māḥið	*mīčil / mīḥið*	*mākil / māḥið*	*ʔākel / ʔāḥed*

Finally, in a few places intensive mutual contact has resulted in an interdialect (Trudgill 1986: 62) with completely new forms, such as the 3PL.M inflectional suffix *-a* in the Syrian village of Ṣōrān (Behnstedt 1994b: 423–425), as shown in Table 4.

Table 4: 3PL.M inflectional suffixes – 'they said'

Bedouin	Sedentary	Ṣōrān
gāḷ-am	*qāl-o*	*qāl-a*

3.3 Syntax

3.3.1 Changes due to contact with Aramaic

3.3.1.1 Clitic doubling

In all but the Bedouin-type dialects of the region, two constructions exist which both use an anticipatory pronoun and the preposition *l-* 'to': (i) a construction

involving analytical marking of a definite direct object, as in (21)–(23); and (ii) a construction involving analytic attribution of a noun, as in (24). The frequency and constraints of these two cases of clitic doubling show great variety, but in general the usage of construction (i) is restricted to specific objects, particularly elements denoting human beings, and construction (ii) is mostly found with inalienable possession, particularly kinship. A detailed discussion of both features is found in Souag (2017).

(21) Damascus (Berlinches 2016: 144)
ḥabbēt-o la-ʕamər
love.PRF.1SG-3SG.M to-Amr
'I loved Amr.'

(22) Baghdad, Christian (Abu-Haidar 1991: 116)
qaɣētū-nu l-əl-əktēb
read.PRF.1SG-3SG.M to-DEF-book
'I read the book.'

(23) Cilician Arabic (*ʕalā* instead of *l-*; Procházka 2002b: 158
biyḥibb-u ʕala ḫāl-u
love.IMPF.IND.3SG.M-3SG.M on uncle-3SG.M
'He loves his (maternal) uncle.'

(24) Baghdad, Christian (Abu-Haidar 1991: 116)
maɣt-u l-aḫū-yi
wife-3SG.M to-brother-OBL.1SG
'my brother's wife'

Though the preposition *l-* is sometimes attested in Classical Arabic for introducing direct objects and is common even in Modern Standard Arabic for analytic noun annexation, there are good arguments that the two constructions are pattern replications of an Aramaic model.[31] For one thing, they do not have direct

[31]Not discussed here are two variants of construction (i), one without the suffix and the other without the preposition (cf. Lentin 2018: 203). Among the many studies that are in favor of Aramaic influence are Contini (1999: 105); Blanc (1964: 130); and Weninger (2011: 750). Diem (1979: 47–49) and Lentin (2018) are more skeptical. Souag (2017: 52) suggests that at least "the initial stages of the development of clitic doubling in the Levant derive from Aramaic substratum influence, but the current situation also reflects subsequent Arabic-internal developments".

parallels either in OA or in dialects which lacked contact with Aramaic. Example (25) shows that both constructions have striking parallels in especially the later eastern varieties of Aramaic (Rubin 2005: 94–104).

(25) a. Syriac (Rubin 2005: 100)
bnā-y l-bayt-ā
build.PRF.3SG.M-3SG.M to-house-DEF
'He built the house.'

b. Syriac (Hopkins 1997: 29)[32]
šm-ēh l-gabr-ā
name-3SG.M to-man-DEF
'the name of the man'

3.3.1.2 *fī* 'can'

In the entire western part of the region including southern Turkey, the preposition *fī* 'in', together with a pronominal suffix, is used to express a capability, as in (26). This has a striking parallel in the modern Aramaic *ʔīθ b-* 'there is in' ~ 'be able' (Borg 2004: 52).

(26) Damascus (Cowell 1964: 415)
fī-ni sāʕd-ak əb-kamm lēra
in-1SG help.IMPF.1SG-2SG.M with-some pound
'Can I help you with a few pounds?'

3.3.1.3 Specific indefinite *šī*

A final example of possible Aramaic influence is the Syrian particle *šī* that mainly indicates partial specifity, as in (27). It might be a pattern replication of the Western Neo-Aramaic form *mett*, used with the same function (Diem 1979: 49). What reduces the likelihood of imposition by Aramaic speakers is the existence of a cognate in Moroccan Arabic which is used with almost the same function.[33]

(27) Damascus (own data)
hnīk fī šī ʕamūd
there EXS INDF column
'There is some column.'

[32]The same pattern using the linker *d-* is more common.
[33]Cf. Brustad (2000: 19, 26–27) and Wilmsen (2014: 51–53).

3.3.2 Changes due to contact with other languages

3.3.2.1 Definiteness

A hallmark of both sedentary and Bedouin-type Iraqi dialects is that reflexes of the noun *fard* 'individual (thing or person)' are used to mark different kinds of indefiniteness (Blanc 1964: 118–119). The same form with the same indefinite article function is found in in the Iranian province of Khuzestan, and in all Arabic speaking language islands of Central Asia, i.e. Khorasan, Uzbekistan, and Afghanistan, as illustrated in (28).

(28) Kirkuk (own data)
taʕrif-lak fadd ṭabīb bāṭiniyye
know.IMPF.2SG.M-DAT.2SG.M INDF doctor internal
'Do you know a doctor of internal medicine?'

It is very likely that the noun *fard* has developed into a kind of indefinite article under the influence of other areal languages, particularly Turkish, Turkmen, Persian, and Neo-Aramaic. However, in contrast to all contact languages, Iraqi Arabic has not grammaticalized the numeral 'one' (*wāḥəd*), but *fard*. This clearly indicates that this feature is a case of pattern replication. There are many parallels in the functions of the indefinite articles (such as marking pragmatic salience, semantic individualization, approximation with numerals). Moreover, in all languages they are not fully systematized as a grammatical category as their usage is often optional.

In the dialects of the Jews of Kurdistan the definite article is often omitted in subject position – a flagrant imitation of the Kurdish model (see also Akkuş, this volume, for some Anatolian dialects).

(29) Kurdistan Arabic (Jastrow 1990: 71)
baʕdēn mudīra baʕatət ḫalf-na
then director send.PRF.3SG.F after-1PL
'Then the director sent for us.'

3.3.2.2 *m-bōr* 'because, in order to'

An interesting case of calquing which shows the difficulty of distinguishing between borrowing and imposition (see Manfredi, this volume) is the conjunction *m-bōr* 'because, in order to'. It exhibits both matter and pattern transfer, as it is a copy of Kurdish *ji ber (ku)*. In the actual form the Kurdish *ji* 'from' was replaced by the Arabic equivalent *m-* (Jastrow 1979: 64).

3.3.2.3 Evidentiality

Syntactic change because of contact with Turkish is restricted to the Arabic dialects spoken in Turkey. In Cilicia and the Harran–Urfa region, active participles express evidentiality, that is, they are used in utterances where a speaker refers to second-hand information. As evidentiality is not a common category in Semitic, it is very likely that the bilingual Arabic speakers of those regions copied this linguistic category from Turkish. In Turkish, any second-hand information is obligatorily marked by the verbal suffix *-mış*, whose second function besides evidentiality is to express stativity and perfectivity. The latter two functions are assumed by the active participle in many Arabic dialects, including those in question here. Thus, we can suppose that the stative/perfective function, which is shared by both Arabic active participles and the Turkish suffix *-mış*, was likely the starting point of the development that led to the additional evidential function of Arabic participles. The fact that evidentials seem to spread readily through language contact (Aikhenvald 2004: 10) makes Turkish influence even more probable.[34] The example in (30) illustrates how the speaker uses perfect forms for those parts of the narrative he witnessed himself, and participles for secondhand information (perfect forms italic, participles in bold face).

(30) Harran–Urfa (Procházka & Batan 2016: 465)
ʔiḥne b-zimānāt *čān* ʕid-na ğār b-al-maḥalle huwwa *māt ərtiḥam* əngūl-lu šēḫ məṭar […] nahāṛ rabīʕ-u wāḥad **ʕāzm**-u ʕala stanbūl **rāyiḥ** maʕzūm ʕala stanbul **māḫið** šēḫ məṭar əb-sāgt-u
‘Once we had a neighbor in our quarter. He died; he passed away. We called him Sheikh Məṭar. One day somebody invited his friend to Istanbul. As he was invited he went to Istanbul and he took Sheikh Məṭar with him.’

3.3.2.4 Comparative and superlative

In most Arabic dialects that are spoken in Turkey, comparatives and superlatives may be expressed by means of the Turkish particles *daha* and *en*, respectively, followed by the simplex instead of the elative form of the adjective (cf. Akkuş, this volume). As for comparatives, the use of such constructions is rather restricted, while, at least in Cilician Arabic, they are relatively frequent for the superlative.

(31) Harran–Urfa (own data)

[34]For more examples and further details see Procházka (2002b: 200–201) for Cilicia, and Procházka & Batan (2016: 464–465) for the Bedouin-type dialects in the Harran–Urfa region.

daha zēn ṣārat
more good become.PRF.3SG.F
'It has become better.'

(32) Cilician Arabic (Procházka 2002b: 155)
mīn en zangīl bi-d-dini
who SUP rich in-DEF-world
'Who is the richest (person) in the world?'

In Cilicia, comparison is often expressed by the elative pattern of an adjective, which is preceded by the particle *issa*. This clearly reflects a calque: the Turkish equivalent of the adverb *issa* 'still, yet' is *daha*, which in Turkish is also used as the particle of the comparative.

(33) Cilician Arabic (Procházka 2002b: 202)
ṣāyir issa aḥsan
become.PTCP more good.ELA
'It became better'.

(34) Turkish
Daha iyi ol-du.
more good become-PRF.3SG
'It became better.'

3.3.2.5 Valency

Sometimes a change in verb valency occurs as a consequence of the copying of Turkish models. A case found throughout these dialects is the verb *ʕağab* 'to like': usually in Arabic the entity that is liked is the grammatical subject and the person who likes something is the direct object of the verb; but in the Arabic dialects in question, the construction of this verb reflects its Turkish (and English) usage with the person doing the liking being the grammatical subject.

(35) a. Cilicia (Procházka 2002b: 200)
ʕğabt bayt-ak
like.PRF.1SG house-2SG.M
b. Damascus (own data)
bēt-ak ʕažab-ni
house-2SG.M like.PRF.3SG.M-1SG
'I liked your house.'

3.4 Lexicon

Apart from the Aramaic loanwords also found in Classical Arabic (see Retsö 2011; van Putten, this volume) – often in the realms of religion and cult – the dialects of this region exhibit a large number of Aramaic lexemes. They are particularly common in Lebanon and western Syria, but also found in Iraq and even in the Bedouin-type dialects (Féghali 1918; Borg 2004; 2008). A large percentage of these words belong to flora and fauna, agriculture, architecture, tools, kitchen utensils, and other material objects:[35]

(36) *ṣumd* ~ *ṣimd* 'plough' < Syriac *ṣāmdē* 'yoke'
qālūz 'bolt (of a door)' < Syriac *qālūzā*
nāṭūr 'guard (of a vineyard etc.)' < Syriac *nāṭūrā*
šaṭaḥ 'to spread' < Syriac *šeṭaḥ*
šōb 'heat, hot' < Syriac *šawbā*

Many nautical terms and words denoting agricultural products and tools were borrowed by Arabic from Greek, often via other languages, especially Aramaic,[36] the Lingua Franca, and Turkish:

(37) *brāṣa* < Greek *práson* 'leek'
laḫana < Greek *láḫana* 'cabbage'
dərrāʔen < Greek *dōrákinon* 'peaches'
ʔabrīm/brīm 'keel' < Greek *prýmnē* 'stern, poop'
sfīn < Greek *sfēn* 'wedge'

Kurdish borrowings are mainly restricted to northern Iraq, where bilingualism is widespread:

(38) Mosul
pūš 'chaff' < Kurdish *pûş*
hēdi hēdi 'slowly' < Kurdish *hêdî* (Jastrow 1979: 68)

The intensive cultural and economic contacts between Iraq and Iran led to many Persian loanwords in various domains of the Iraqi dialects.

[35] See also Neishtadt (2015: 282). Note that, unless otherwise indicated, lexemes cited in this section are taken from Barthélemy (1935) for Syrian dialects, and Woodhead & Beene (1967) and al-Bakrī (1972) for Iraqi dialects.

[36] This is especially true for words related to Christian liturgy and ritual, which constitute about twenty per cent of the Greek vocabulary that entered the dialects of Syria.

(39) *mēwa* 'fruit' < Persian *mīva ~ mayva*
baḫat 'luck' < Persian *baḫt*
čariḫ 'wheel' < Persian *čarḫ*
gulguli 'pink' < Persian *gol* 'rose'
yawāš 'slow' < Persian *yavāš*
puḫta 'mush' < Persian *poḫte* '(well) cooked'

Ottoman Turkish contributed a great deal to culinary vocabulary and the terminology of clothing and (technical) tools of Syria and Iraq.[37] It was even the source of several adverbs and even verbs in the local Arabic varieties (Halasi-Kun 1969; 1973; 1982).

(40) Syria (Damascus)
šāwərma 'shawarma' < Turkish *çevirme*
ṣāž 'iron plate for making bread' < Turkish *saç*
yalanži 'vine-leaves stuffed with rice' < Turkish *yalancı* 'liar' (as they pretend to be "real" *dolma* stuffed with meat)
šīš ṭāwū? 'spit-roasted chicken' < Turkish *şiş tavuk*
kəzlok 'glasses' < Turkish *gözlük*
ʔūḍa 'room' < Turkish *oda*
ballaš 'to begin' < Turkish *başla-mak* by metathesis.

(41) Iraq (Muslim Baghdadi, cf. Reinkowski 1995)
qūzi 'a dish with roasted mutton' < Turkish *kuzu* 'lamb'
tēl 'wire' < Turkish *tel*
yašmāɣ 'kerchief (for men)' < Turkish *yaşmak* 'veil (for women)'
bōš 'empty; neutral', which yielded also the verb *bawwaš* 'to put into neutral (gear)' < Turkish *boş* 'empty'
qačaɣ 'smuggled goods' < Turkish *kaçak*

During the last century, the Arabic dialects in Turkey[38] have incorporated numerous Turkish words in addition to loanwords from Ottoman times. Among them are terms in education, medicine, sports, media, and technology. Besides these, kinship terms, the vocabulary of everyday life, and structural words like adverbs and discourse markers have infiltrated the dialects from Turkish.

[37] The same loanwords are, of course, often found in other regions that were under Ottoman rule, above all in Egypt, but also in Tunisia, Yemen and other regions.

[38] For Cilicia see Procházka (2002b; 2002a: 187–199).

(42) Cilician Arabic
qāyin … '-in-law' (< Turkish *kayın*)
ṭōrūn 'grandchild' (< Turkish *torun*)
bīle 'even' (< Turkish *bile*)
qāršīt 'opposite from' (< Turkish *karşı*)

The cases of semantic extension of an Arabic word result from the wider semantic range of its Turkish equivalent which has been transferred into Arabic. Thus, in both Cilician and Harran–Urfa Arabic *sāq/ysūq* 'to drive' also occurs with the meaning of 'to last' like the Turkish verb *sürmek*. In Harran–Urfa *b-arḍ* 'on the place/ground (of)' has become a preposition/conjunction meaning 'instead'. This can be seen as an instance of contact-induced grammaticalization (Gardani et al. 2015: 4) under the influence of Turkish *yerine* 'instead, in its place'.

(43) Harran–Urfa (own data)
al-mille tākl-u b-arḍ al-laḥam
DEF-people eat.IMPF.3SG.F-3SG.M in-place DEF-meat
'The people eat it instead of meat.'

(44) Harran–Urfa (own data)
b-arḍ-in tibči ʔigir āya
in-place-LINK cry.IMPF.2SG.M read.IMP.SG.M verse
'Instead of crying recite a (Koranic) verse!'

In Iraq, many English words related to Western culture and technology have been, and still are, borrowed into the dialects. The same is true for French in Syria and (particularly) Lebanon (cf. Barbot 1961: 176).

(45) Iraq (words of English origin)
kitli < kettle
buṭil < bottle
glāṣ < glass
pančar 'flat tire' (< *puncture*)
pāysikil < bicycle
māṭōrsikil < motorcycle
lōri < lorry
igzōz < exhaust (pipe)
brēk < brake

(46) Syria and Lebanon (words of French origin)
gātto ~ gaṭō < *gâteau* 'cake'
garsōn < *garçon* 'waiter'
sēšwār < *séchoir* 'hair drier'
kwaffēr < *coiffeur* 'hair-dresser'
ʔaṣanṣēr < *ascenseur* 'elevator'
grīb < *grippe* 'influenza'

Due to long-term contacts, there are mutual borrowings between the Bedouin and sedentary dialects of the region. This affects not only specific vocabulary of respective culture but also basic lexical items. Historically, the sedentary dialects have been much more influenced by the Bedouin-type dialects than vice versa.

4 Conclusion

The sociolinguistic history of the regions treated here suggests that the conditions for imposition were relatively restrictive and mainly found in contact settings with Aramaic which, over the centuries, has been given up by most of its speakers in favor of Arabic. Thus, it is not surprising that so many features beyond the lexicon for which contact-induced change can be assumed are related to Aramaic influence.

Morphological borrowing is in general relatively rare because it presupposes a high intensity of contact (Gardani et al. 2015: 1). Practically all cases presented in §2.2 corroborate the universal tendencies that: (i) derivational morphology is more prone to borrowing than inflectional morphology; and (ii) nominalizers and diminutives are very frequently represented in instances of borrowed derivational morphology (Gardani et al. 2015: 7; Seifart 2013). On the whole, the Bedouin-type dialects exhibit significantly fewer contact-induced changes than the sedentary dialects. This may be the result of both the Bedouin groups' nomadic way of life at the fringes of the desert and their tribally organized society, which impedes intense contact with outsiders.

The relative infrequency of contact-induced changes in morphology and syntax found in the Arabic varieties spoken in Turkey have two main reasons: first, the high degree of complete bilingualism is a very recent phenomenon that only pertains to the last two generations; and second, and probably more importantly, the great structural differences between the two languages, which have impeded both matter and pattern replications.

What is still relatively unclear is the degree of historical bilingualism between Arabic on the one hand and Ottoman Turkish, Kurdish, and Persian on the other. Future research would be particularly desirable with regard to Iraq, providing interesting new data on contact-induced changes in multilingual regions like Mosul and Kirkuk where Arabic, Turkmen, and Kurdish speakers have been in contact for a long time. Also, studies like that of Neishtadt (2015) for Palestine should be carried out for Syrian and especially Iraqi dialects with regard to lexical borrowings from Aramaic. Another completely under-researched topic is idiomatic constructions, in which the mutual influence of most languages in the region may be assumed.

Further reading

There are no studies which treat the subject of contacts between Arabic and the other languages of the whole region covered in this chapter. However:
Arnold & Behnstedt (1993) is an in-depth study of the mutual contacts between Western Neo-Aramaic and the local Arabic dialects in the Anti-Lebanon Mountains of Syria.
Diem (1979) is a pioneer study of substrate influence in the modern Arabic dialects, though with focus on South Arabia, i.e. outside of the region treated in this chapter.
Palva (2009) is a very good case study of the diachronic relations between sedentary and Bedouin-type dialects in the Iraqi capital Baghdad.
Weninger (2011) is a concise overview of contact between different varieties of Aramaic and Arabic.

Acknowledgements

I am grateful to my colleagues Bettina Leitner and Veronika Ritt-Benmimoun for their valuable comments on earlier drafts of this paper. I warmly thank Jérôme Lentin for extensive discussion of the possible origin of the hypocoristic *-o* suffix (§3.2.4) and his help in finding important sources.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person
BCE	before Common Era
CE	Common Era
COMP	complementizer
DEF	definite
F	feminine
ELA	elative degree
EXS	existential
IMP	imperative
IMPF	imperfect (prefix conjugation)
INDF	indefinite
L1	first language
L2	second language
LINK	linker
M	masculine
OA	Old Arabic
OBL	oblique
PL	plural
PRF	perfect (suffix conjugation)
SG	singular
SUP	superlative

References

Abu-Haidar, Farida. 1991. *Christian Arabic of Baghdad*. Wiesbaden: Harrassowitz.

Abu-Haidar, Farida. 1999. Vocatives as exclamatory nouns in Iraqi Arabic. In Yasir Suleiman (ed.), *Arabic grammar and linguistics*, 144–160. London: Routledge.

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.

al-Bakrī, Ḥāzim. 1972. *Dirāsāt fī l-alfāḍ al-ʕāmiyya al-mawṣiliyya*. Baghdad: Maṭbaʕat Asʕad.

Arnold, Werner. 1998. *Die arabischen Dialekte Antiochiens*. Wiesbaden: Harrassowitz.

Arnold, Werner & Peter Behnstedt. 1993. *Arabisch-aramäische Sprachbeziehungen im Qalamūn (Syrien)*. Wiesbaden: Harrassowitz.

Barbot, Michel. 1961. Emprunts et phonologie dans les dialectes citadins syro-libanais. *Arabica* 8(2). 174–188.

Barthélemy, A. 1935. *Dictionnaire arabe-français, dialectes de Syrie: Alep, Damas, Liban, Jérusalem*. Vol. 5. Paris: Paul Geuthner. Volumes 4–5 edited by H. Fleisch.

Behnstedt, Peter. 1992. Zur Spaltung von *ā in *ō/ē* (*o̯/e̯*) im Qalamūn (Antilibanon). *Die Welt des Orients* 23. 77–100.

Behnstedt, Peter. 1994a. *Der arabische Dialekt von Soukhne (Syrien). Teil 2: Phonologie, Morphologie, Syntax. Teil 3: Glossar*. Wiesbaden: Harrassowitz.

Behnstedt, Peter. 1994b. Dialektkontakt in Syrien. In Dominique Caubet & Martine Vanhove (eds.), *Actes des premières journées internationales de dialectologie arabe de Paris*, 417–430. Paris: Langues'O.

Behnstedt, Peter. 1997. *Sprachatlas von Syrien*. Vol. I: Kartenband. Wiesbaden: Harrassowitz.

Berlinches, Carmen. 2016. *El dialecto árabe de Damasco (Siria): Estudio gramatical y textos*. Zaragoza: Prensas de la Universidad de Zaragoza.

Blanc, Haim. 1964. *Communal dialects in Baghdad*. Cambridge, MA: Harvard University Press.

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. 2004. *A comparative glossary of Cypriot Maronite Arabic (Arabic–English): With an introductory essay*. Leiden: Brill.

Borg, Alexander. 2008. Lexical Aramaisms in nomadic Arabic vernaculars. In Stephan Procházka & Veronika Ritt-Benmimoun (eds.), *Between the Atlantic and Indian Oceans: Studies on contemporary Arabic dialects. Proceedings of the 7th AIDA Conference, held in Vienna from 5-9 September 2006*, 89–112. Münster: LIT-Verlag.

Brockelmann, Carl. 1928. Deminutiv und Augmentativ im Semitischen. *Zeitschrift für Semitistik und verwandte Gebiete* 6. 109–134.

Brustad, Kristen. 2000. *The syntax of spoken Arabic: A comparative study of Moroccan, Egyptian, Syrian, and Kuwaiti dialects*. Georgetown: Georgetown University Press.

Cantineau, Jean. 1960. *Etudes de linguistique arabe*. Paris: Klincksieck.

Coghill, Eleanor. 2012. Parallels in the grammaticalisation of Neo-aramaic *zil-* and Arabic *raḥ-* and a possible contact scenario. In Domenyk Eades (ed.), *Grammaticalization in Semitic*, 127–144. Oxford: Oxford University Press.

Contini, Ricardo. 1999. Le substrat araméen en néo-arabe libanais: Préliminaires à une enquête systématique. In Marcello Lamberti & Livia Tonelli (eds.), *Afro-asiatica Tergestina*, 101–122. Padova: Unipress.

Corriente, Federico. 1969. *Qalqūl* en semitico. *Sefarad* 29. 3–11.

Cowell, Mark W. 1964. *A reference grammar of Syrian Arabic*. Washington, DC: Georgetown University Press.

Diem, Werner. 1979. Studien zur Frage des Substrats im Arabischen. *Der Islam* 56. 12–80.

Donner, Fred McGraw. 1981. *The early Islamic conquests*. Princeton: Princeton University Press.

Fassberg, Steven Ellis. 2010. *The Jewish Neo-Aramaic dialect of Challa*. Leiden: Brill.

Féghali, Michel. 1918. *Etude sur les emprunts syriaques dans les parlers arabes du Liban*. Paris: Champion.

Ferguson, Charles. 1969. The /g/ in Syrian Arabic: Filling a gap in a phonological pattern. *Word* 25. 114–119.

Ferguson, Charles A. 1997. Arabic baby talk. In Kirk Belnap & Niloofar Haeri (eds.), *Structuralist studies in Arabic linguistics: Charles A. Ferguson's papers, 1954–1994*, 179–187. Leiden: Brill.

Fleisch, Henri. 1974a. Le changement *ā* > *ō* dans le sémitique de l'ouest et en arabe dialectal libanais. In Henri Fleisch (ed.), *Etudes d'arabe dialectal*, 45–50. Beirut: Dar el-Machreq.

Fleisch, Henri. 1974b. Le parler arabe de Kfar-Ṣghâb, Liban. In Henri Fleisch (ed.), *Etudes d'arabe dialectal*, 221–262. Beirut: Dar el-Machreq.

Gardani, Francesco, Peter Arkadiev & Nino Amiridze (eds.). 2015. *Borrowed morphology*. Berlin: De Gruyter.

Gazsi, Dénes. 2011. Arabic–Persian language contact. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1015–1021. Berlin: De Gruyter Mouton.

Gralla, Sabine. 2006. *Der arabische Dialekt von Nabk (Syrien)*. Wiesbaden: Harrassowitz.

Grigore, George. 2007. *L'arabe parlé à Mardin: Monographie d'un parler arabe "périphérique"*. Bucharest: Editura Universitãţii din Bucureşti.

Halasi-Kun, Tibor. 1969. The Ottoman elements in the Syrian dialects. *Archivum Ottomanicum* 1. 14–91.

Halasi-Kun, Tibor. 1973. The Ottoman elements in the Syrian dialects, II. *Archivum Ottomanicum* 5. 17–95.

Halasi-Kun, Tibor. 1982. The Ottoman elements in the Syrian dialects, II (continuation). *Archivum Ottomanicum* 7. 117–267.

Holes, Clive. 2002. Non-Arabic Semitic elements in the Arabic dialects of eastern Arabia. 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*, 269–279. Wiesbaden: Harrassowitz.

Holes, Clive. 2007. Colloquial Iraqi Arabic. In J. N. Postgate (ed.), *Languages of Iraq, ancient and modern*, 123–134. Baghdad: British School of Archaeology in Iraq.

Holes, Clive. 2016. *Dialect, culture, and society in eastern Arabia*. Vol. 3: *Phonology, morphology, syntax, style*. Leiden: Brill.

Hopkins, Simon. 1995. *ṣarār* 'pebbles': A Canaanite substrate word in Palestinian Arabic. *Zeitschrift für Arabische Linguistik* 30. 37–49.

Hopkins, Simon. 1997. On the construction *šmēh l-gabrā* 'the name of the man' in Aramaic. *Journal of Semitic Studies* 42. 23–32.

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen* qəltu*-Dialekte*. Vol. 1: *Phonologie und Morphologie*. Wiesbaden: Steiner.

Jastrow, Otto. 1979. Zur arabischen Mundart von Mossul. *Zeitschrift für Arabische Linguistik* 2. 76–99.

Jastrow, Otto. 1990. Die arabischen Dialekte der irakischen Juden. In Werner Diem & Abdoldjavad Falaturi (eds.), *XXIV. Deutscher Orientalistentag vom 26. bis 30. September 1988 in Köln: Ausgewählte Vorträge*, 199–205. Wiesbaden: Steiner.

Kahane, Henry Romanos, Renée Kahane & Andreas Tietze. 1958. *The Lingua Franca in the Levant: Turkish nautical terms of Italian and Greek origin*. Urbana, IL: University of Illinois Press.

Khoury, Philipp S. 1987. *Syria and the French Mandate: The politics of Arab nationalism, 1920–1945*. London: Tauris.

Krebernik, Manfred. 2008. Von Gindibu bis Muḥammad: Stand, Probleme und Aufgaben altorientalistisch-arabistischer Philologie. In Otto Jastrow, Shabo Talay & Herta Hafenrichter (eds.), *Studien zur Semitistik und Arabistik: Festschrift für Hartmut Bobzin zum 60. Geburtstag*, 247–279. Wiesbaden: Harrassowitz.

Lentin, Jérôme. 2018. The Levant. In Clive Holes (ed.), *Arabic historical dialectology: Linguistic and sociolinguistic approaches*, 170–205. Oxford: Oxford University Press.

Lucas, Christopher. 2015. Contact-induced language change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.

Magidow, Alexander. 2013. *Towards a sociohistorical reconstruction of pre-Islamic Arabic dialect diversity*. Austin: University of Texas. (Doctoral dissertation).

Masliyah, Sadok. 1996. Four turkish suffixes in Iraqi Arabic: *-li*, *-lik*, *-siz* and *-çi*. *Journal of Semitic Studies* 41(2). 291–300.

Masliyah, Sadok. 1997. The diminutive in spoken Iraqi Arabic. *Zeitschrift für Arabische Linguistik* 33. 68–88.

Muraoka, Takamitsu. 2005. *Classical Syriac*. 2nd, revised edn. Wiesbaden: Harrassowitz.

Neishtadt, Mila. 2015. The lexical component in the Aramaic substrate of Palestinian Arabic. In Aaron Butts (ed.), *Semitic languages in contact*, 280–310. Leiden: Brill.

Nöldeke, Theodor. 1904. *Neue Beiträge zur semitischen Sprachwissenschaft*. Strasbourg: Trübner.

Owens, Jonathan. 2006. *A linguistic history of Arabic*. Oxford: Oxford University Press.

Palva, Heikki. 2009. From *qəltu* to *gilit*: Diachronic notes on linguistic adaptation in Muslim Baghdad Arabic. In Enam Al-Wer & Rudolf De Jong (eds.), *Arabic dialectology in honour of Clive Holes on the occasion of his sixtieth birthday*, 17–40. Leiden: Brill.

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.

Perry, John R. 2007. Persian morphology. In Alan S. Kaye (ed.), *Morphologies of Asia and Africa*, 975–1019. Winona Lake, IN: Eisenbrauns.

Procházka-Eisl, Gisela. 2018. A suffix on the move: Forms and functions of the Turkish suffix *-ci* in Arabic dialects. *Mediterranean Language Review* 25. 21–52.

Procházka, Stephan. 2002a. Contact phenomena, code-copying, and code-switching in the Arabic dialects of Adana and Mersin (southern Turkey). In Abderrahim Youssi (ed.), *Aspects of the dialects of Arabic today: Proceedings of the 4th conference of the International Arabic Dialectology Association (AIDA), Marrakesh, April 1-4, 2000. In honour of Professor David Cohen*, 133–139. Rabat: Amapatril.

Procházka, Stephan. 2002b. *Die arabischen Dialekte der ̧Cukurova (Südtürkei)*. Wiesbaden.

Procházka, Stephan. 2004. Unmarked feminine nouns in modern Arabic dialects. In Martine Haak, Rudolf de Jong & Kees Versteegh (eds.), *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, 237–282. Leiden: Brill.

Procházka, Stephan. 2013. Traditional boatbuilding: Two texts in the Arabic dialect of the island of Arwād (Syria). In Renaud Kuty, Ulrich Seeger & Shabo Talay (eds.), *Nicht nur mit Engelszungen: Beiträge zur semitischen Dialektologie. Festschrift für Werner Arnold zum 60. Geburtstag*, 275–288. Wiesbaden: Harrassowitz.

Procházka, Stephan. 2018. The northern Fertile Crescent. In Clive Holes (ed.), *Arabic historical dialectology: Linguistic and sociolinguistic approaches*, 257–292. Oxford: Oxford University Press.

Procházka, Stephan & Ismail Batan. 2016. The functions of active participles in Šāwi bedouin dialects. In George Grigore & Gabriel Biţună (eds.), *Arabic va-*

rieties: Far and wide. Proceedings of the 11th International Conference of Aida, Bucharest 2015, 457–466. Bucharest: Editura Universităţii din Bucureşti.

Prunet, Jean-François & Ali Idrissi. 2014. Overlapping morphologies in Arabic hypocoristics. In Sabrina Bendjaballah, Noam Faust, Mohamed Lahrouchi & Nicola Lampitelli (eds.), *The form of structure, the structure of form: Essays in honor of Jean Lowenstamm*, 177–192. Amsterdam: John Benjamins.

Reinkowski, Maurus. 1995. Eine phonologische Analyse türkischen Wortgutes im Bagdadisch-Arabischen mitsamt einer Wortliste. *Folia Orientalia* 31. 89–115.

Retsö, Jan. 2011. Aramaic/Syriac loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Rubin, Aaron D. 2005. *Studies in Semitic grammaticalization*. Winona Lake, IN: Eisenbrauns.

Sabuni, Abdulghafur. 1980. *Laut- und Formenlehre des arabischen Dialekts von Aleppo*. Frankfurt am Main: Peter Lang.

Seifart, Frank. 2013. *AfBo: A world-wide survey of affix borrowing*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://afbo.info.

Shields, Sarah. 2004. Mosul questions: Economy, identity, and annexation. In Reeva S. Simon & Eleanor Harvey Tejirian (eds.), *The creation of Iraq: 1914–1921*, 50–60. New York: Columbia University Press.

Singer, Hans-Rudolf. 1984. *Grammatik der arabischen Mundart der Medina von Tunis*. Berlin: De Gruyter.

Souag, Lameen. 2017. Clitic doubling and language contact in Arabic. *Zeitschrift für Arabische Linguistik* 66. 45–70.

Talay, Shabo. 1999. *Der arabische Dialekt der Khawētna*. Vol. 1: Grammatik. Wiesbaden: Harrassowitz.

Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Wiley-Blackwell.

Weninger, Stefan. 2011. Aramaic–Arabic language contact. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 747–755. Berlin: De Gruyter Mouton.

Wilkins, Charles L. 2010. *Forging urban solidarities: Ottoman Aleppo 1640–1700*. Leiden: Brill.

Wilmsen, David. 2014. *Arabic indefinites, interrogatives and negators: A linguistic history of western dialects*. Oxford: Oxford University Press.

Woidich, Manfred. 2006. *Das Kairenisch-Arabische: Eine Grammatik*. Wiesbaden: Harrassowitz.

Woodhead, Daniel R & Wayne Beene. 1967. *A dictionary of Iraqi Arabic: Arabic–English*. Washington, DC: Georgetown University Press.

Chapter 4

Andalusi Arabic

Ángeles Vicente
University of Zaragoza

This chapter covers an ancient contact language situation: Andalusi Arabic with another two languages – the Romance varieties spoken by the local population, and the Berber varieties brought by different Berber speakers arriving in al-Andalus during its existence. The situation of bilingualism whereby the Romance language was sociolinguistically dominant for most of the population over the course of several centuries resulted in numerous contact-induced changes in all areas of grammar. In addition, interaction between Arabic-speaking and Berber-speaking populations constituted a second locus of language contact with consequences for Andalusi Arabic.

1 Historical development of Andalusi Arabic

A dialect of the western Neo-Arabic type, Andalusi Arabic is currently a dead language. It was spoken from the eighth to the seventeenth century in a changing territory following historical vicissitudes.

Arabic arrived in the Iberian Peninsula in the eighth century with Arabic-speaking tribes coming from different zones at various stages.[1] According to historical sources, the number of Muslims initially arriving was small, most of them probably partially Arabized Berber-speakers from North Africa.[2] Over time, the

[1]Historians have long argued for the ethnic variety of the Arabs who invaded the Iberian Peninsula, particularly referring to the presence of Syrian and Yemeni tribes. See Terés Sádaba (1957), Al-Wasif (1990) and Guichard (1995).

[2]Historians agree that it is extremely difficult, if not impossible, to establish what the level of Arabization of this population was. According to Manzano Moreno (1990: 399), it seems that linguistic Arabization was not widespread among Andalusi Berbers at least during the eighth century.

Ángeles Vicente. 2020. Andalusi Arabic. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 83–102. Berlin: Language Science Press. DOI:??

society of al-Andalus (the name given to the territory in the Iberian Peninsula under different Muslim–Arab systems of rule for eight centuries) would eventually come to use a distinctive variety of Maghrebi Arabic known as Andalusi Arabic.[3] This variety evolved through dialectal levelling and changes resulting from contact with other languages present in the zone, and became a reasonably unified variety by the tenth century. The political success of the Umayyad dynasty and the establishment of their caliphate in the year 929 CE may have contributed to language levelling, though dialect variation continued to exist in the form of diatopical variants from various regions; scholars thus refer to the existence of an Andalusi "dialect bundle" (e.g. Corriente 1977: 6; 1992b: 446). For instance, the Granadian variety seems to have been more conservative than those spoken in other regions.[4] The regional Andalusi variety spoken in Valencia was the last to disappear with the expulsion of the *moriscos* (Muslims forced to convert to Christianity) in the seventeenth century (Barceló & Labarta 2009: 117).

Even though Andalusi Arabic was a vernacular variety, the few extant sources are always written, and therefore reflect a higher register than that of the language used for daily communication. In fact, hardly any material reflecting the everyday dialectal level is available, since most of the sources consist of texts written in Middle Arabic (i.e. a written form intermediate between Classical and spoken dialectal Arabic; see Lentin 2011). Furthermore, complications arise due to the use of Arabic script to record dialect variants.[5]

Consequently, a comprehensive view of all the periods and places where this language was spoken is lacking. For instance, sources are scarce regarding the use of the language in the eighth and ninth centuries. As Wasserstein (1991: 3) put it: "A linguistic map of Islamic Spain for any period between the middle of the eighth century and the middle of the thirteenth century would be extremely difficult to draw."

Notwithstanding this, written documents in Andalusi Arabic are available from the tenth century until the expulsion of the *moriscos* in the seventeenth century.

[3] Andalusi Arabic features the only common discriminating trait of Maghrebi varieties, that is, the *n-* and *n-…-u* desinences for the first person singular and plural of the imperfect (cf. Benkato, this volume).

[4] According to Corriente (1998b: 56), this is because Granada was relatively isolated from the Andalusi mainstream, and played a secondary political role, at least initially. An example that Corriente gives of this conservatism is the retention of strong *imāla* (raising of originally low front vowels) found in Granadian Arabic, since this feature was eliminated or reduced in other Andalusi varieties with written attestation.

[5] An overview of sources for the description of Andalusi Arabic can be found in Corriente et al. (2015: xxiii–xxiv).

The oldest documented and preserved Andalusi text is an early form of *zaǧal* poetry dating from 913 CE, illustrated in (1).[6]

(1) Tenth-century Andalusi Arabic (Corriente et al. 2015: 237).[7]

a. labán úmm-u fi fúmm-u
milk mother-3SG.M in mouth-3SG.M
'His mother's milk is in his mouth.'

b. rás ban ḥafṣún fi ḥúkm-u
head Ban Ḥafṣún in power-3SG.M
'Ban Ḥafṣún's head is at his disposal.'

The latest attestations of this language consist of private documents written by *moriscos* from Valencia from seventeenth century, in which interesting instances of Romance dialectalisms and influence of Catalan, the Romance language spoken in the region, Aragonese and Castilian can be observed (Barceló & Labarta 2009: 119).

Andalusi Arabic continued to be spoken in the Iberian Peninsula after the end of al-Andalus as a Muslim–Arab state in 1492 CE, as some of the Arabic-speaking population remained in certain regions up until the seventeenth century, when the last *moriscos* were expelled. This language was therefore taken by the migrant population to various places in North Africa in different periods from the Middle Ages up to the Modern Era.[8]

Initially the second language (L2) of most of the population, after a two-century gestation process (approximately from the conquest in 711 to the beginning of the caliphate in 929), Andalusi Arabic gradually became the first language (L1) of the majority of the population, overtaking the Romance dialect spoken by the original local population. The main reason for this was the growing social prestige attached to Arabic in an Islamic society, in contrast to the lower social status of Andalusi Romance, which became the local L2 and eventually disappeared.[9]

Andalusi Arabic became the dominant language (regardless of religion) thanks to the political and social situation of al-Andalus. Furthermore, the advent of an

[6] It consists of a verse by one of the supporters of ʕUmar ibn Ḥafsūn, insulting the caliph ʕAbd ar-Raḥmān III. It appears in the historical chronicle *al-Muqtabis V*, by Ibn Ḥayyān.

[7] Acute accents on vowels in transcription of Andalusi Arabic represent stress rather than vowel length. See §3.1.1 for further details.

[8] This is the reason why Andalusi Arabic has played a very important role in the formation of the Moroccan Arabic (cf. Vicente 2010; Heath, this volume).

[9] Mixed marriages between Muslims and Christian women constituted a significant factor in the propagation of Andalusi Arabic amongst Christians until it also became their L1 (Guichard 1989: 82–83; 1995: 456–457; Chalmeta 2003).

Arabic-speaking population from the east, especially in the Umayyad caliphate (929–1031), played a major role in the expansion of Arabization. According to some scholars such as Fierro Bello (2001) and Corriente (2008: 104), al-Andalus became a society largely monolingual in Andalusi Arabic around the eleventh century, though communities using other languages did exist, especially in rural areas (see §2.1 for more details).

The vernacular Arabic variety spoken in al-Andalus even reached the status of a literary language, appropriating part of the domain of Classical Arabic through proverbs and a number of stanza-based poetic forms (including some *ẖaraǧāt* and the *azǧāl*). Andalusi Arabic poetry reached the circles of the court and the palaces of Taifa Kings. Such social and cultural prestige reveals the extent to which Andalusi Arabic had become the dominant language in this society, and it is for this reason that it is the best-documented vernacular Arabic variety of all those spoken in the Middle Ages.

Andalusi Arabic does not conform neatly to either the Bedouin or the pre-Hilali sedentary type of dialect in the classification usually applied nowdays to Maghrebi Arabic dialects (cf. Benkato, this volume). It shares features of both types of dialects. For instance, in the phonological system, the three interdental phonemes are the same as those in Old Arabic, as is the case in Bedouin-type Maghrebi dialects;[10] however, /q/ is realized using the voiceless variant [q] as in sedentary-type dialects, rather than the voiced variant [g], as in Bedouin-type dialects.[11]

According to Corriente (1992a: 34), the number of speakers of Andalusi Arabic was at its greatest between the eleventh century – a time when the Andalusi koine reached maturity – and the twelfth century.

2 Contact languages

Andalusi Arabic developed in the Iberian Peninsula through the interaction of various Arabic dialects along with two contact languages.[12] This situation spanned

[10]In sedentary-type Maghrebi dialects these are typically pronounced as occlusives. The data do show that the occlusive pronunciation of interdentals was known in Andalusi Arabic, though it was considered vulgar and was repressed (Corriente et al. 2015: 29).

[11]That said, /q/ may have been realized as a voiced [g] in some registers, regions or periods in Andalusi Arabic (see Corriente et al. 2015: 64).

[12]Besides eastern Neo-Arabic varieties brought by invaders in the eighth century, from which Andalusi Arabic emerged, this language continued to evolve in interaction with Maghrebi dialects, particularly with Moroccan Arabic. Owing to this, it is possible to find intra-Arabic contact-induced language change, for instance in the Andalusi variety of Granada. Some instances of trans-

a long period of time, resulting in a significant amount of transfer. This has been analysed by various authors (e.g. Ferrando 1995; 1997; Vicente 2006), and particularly by Corriente (e.g. Corriente 1981; 1992a; 2000; 2002).

The languages with which Andalusi Arabic was in contact were the Romance varieties spoken by the Andalusi population and the Berber varieties brought by different Berber speakers arriving in al-Andalus during its existence.

2.1 Andalusi Romance

Andalusi Romance is a dialect bundle originating in the Romance varieties that were spoken in the Iberian Peninsula when the Islamic invasion occurred in 711,[13] and which underwent a particular evolution through interaction with Arabic. This Ibero-Romance dialect was the L1 of a large proportion of Andalusi society regardless of their religion. It is also the oldest documented variety of Ibero-Romance: according to Corriente, the language of the *ẖarağāt* (see below) reflects the Romance dialect bundle used in al-Andalus between the ninth and eleventh centuries (Corriente 1995; 1997b; 2000).

The language is not well known: only a few written sources are available, transmitted by copyists who may have had limited knowledge of the language. These sources are written both in Arabic and Latin scripts.

Sources in Arabic script consist of bilingual dictionaries and botanical, agronomical and medical glossaries. These evidence a limited number of Andalusi Romance loanwords in Andalusi Arabic, constituting less than 5% of the lexicon according to Corriente (1992a: 142).

Another source in Arabic script are *ẖarağāt*, the final refrains of each stanza of the *muwaššaḥāt*, one of the two types of Andalusi strophic poetry. A few of these refrains were partially written in Andalusi Romance.[14] In addition to these

fer from Moroccan are the verbs *šāf* 'to see' and *ǧāb* 'to bring', and the second element in the negative *ma šāf ši* 'he did not see' (cf. Corriente 1998b: 57). For example, the particle *lás* or *lís* (a variant of *lás* with *imāla*) was the most frequently used negation particle in Andalusi Arabic, while the *ma... ši* construction was generally exceptional in older sources, though not in the work of aš-Šustarī, a Granadian author, due to his travels to North Africa, according to Corriente et al. (2015: 212–215). In addition, Classical Arabic had an influence, especially on the lexicon. The migration of the Bedouin population into North Africa, however, did not have an influence on the evolution of Andalusi Arabic.

[13]These varieties in turn descended from Iberian Vulgar Latin, with substrate influence from pre-Romance Iberian languages and Visigoth lexical borrowings.

[14]Up to 68 *ẖarağāt* in Andalusi Romance have been found (42 in Arabic script and 26 in Hebrew script) with one or more words in this language (Corriente 1997b: 268–323), all of them dating from the tenth–eleventh centuries (Corriente 1997b: 343).

ẖarağāt, loanwords of Andalusi Romance origin were also transmitted in the *zağal* poems of Ibn Quzman.

Latin-script sources also exist, in toponymy, for instance, as well as in loanwords from Andalusi Romance in more northerly Romance dialects, though the data these contribute need to be treated with caution, since adaptation to other Romance dialects blurs features of the source language, making them of limited use from a linguistic point of view.

Andalusi Romance has been analysed by Corriente (1995; 2000; 2012); who has compiled lists of lexical borrowings from Andalusi Romance into Andalusi Arabic in botanical glossaries and in *ẖarağāt* poetry.

In the first centuries of the history of al-Andalus, Andalusi Romance was the L1 used by the majority of Andalusi society, even by some Muslims, such as the *muwalladūn* (converted Muslims), who would learn Arabic as their L2 for self-promotion in society. In time, however, as an Arabic variety became the dominant language, diastratic differences become noticeable. Thus, Andalusi Romance was the L1 used by the rural population and lower classes, whereas the urban Andalusi population underwent more rapid Arabization due to increased exposure to Arabic through mosques, schools, trade, pilgrimages, and so on. Thus, the inhabitants of cities and, above all, leading members of society always had Andalusi Arabic as their L1.

No concrete evidence exists as to when monolingualism in Andalusi Arabic became established. The most commonly accepted date for the disappearance of Romance as a common means of communication in al-Andalus is the late twelfth century, under Almoravid rule. This period saw migrations north out of al-Andalus of the Christian Mozarabs, although most of these were in fact Arabic speakers, as instances of lexical borrowings from Andalusi Arabic in Romance languages from the north reveal. Corriente (1997a; 1992a: 443; 2005) suggests that bilingualism no longer existed by the thirteenth century, and that in the eleventh and twelfth centuries was merely vestigial. In contrast, Galmés de Fuentes and Menéndez Pidal have defended the existence of bilingualism in Andalusi society up until the thirteenth century (Galmés de Fuentes 1994: 81–88; Menéndez Pidal & de Fuentes 2001).[15]

[15]While some Romance-speaking communities may indeed have lasted up until the thirteenth century, note that this circumstance does not imply the existence of a wider bilingual Andalusi society.

2.2 Berber

The arrival of a Berber-speaking population in al-Andalus took place in the eighth and thirteenth centuries, first as auxiliary troops and later as conquerors, though many of them may have already become Arabic-speaking and used an early form of North African Arabic as L2 or even as L1 in the case of those arriving later.

Modern historiography (e.g. Manzano Moreno 1990; Guichard 1995; Chalmeta 2003) reveals that a significant number of Berbers played a major role in the conquest of al-Andalus, a population which grew larger with the later arrival of the Almoravid and Almohad dynasties in the twelfth and thirteenth centuries. Interaction between Arabic-speaking and Berber-speaking populations on both sides of the Strait of Gibraltar facilitated lasting language contact.

The role of Berber in the language development of al-Andalus has not been analysed in depth, however. This is due to data being scarce regarding not only the state of Berber varieties at the time, but also their impact on Andalusi Arabic and the speed of their disappearance from the language scene in the Iberian Peninsula. No sources exist written directly in Berber, plus interpretation issues arise due to the transmission of Berber loanwords in Arabic or Latin script, as the phonological systems do not fully coincide.

Berber varieties had no social prestige in al-Andalus, and were associated with lower registers, a fact which had an obvious repercussion on the direction of transfers in contact-induced changes. According to scholars such as Chalmeta (2003: 160) and Guichard (1995) the reason behind this could be the Berbers' social organization, who tended to settle in rural zones.

As a result of all of the above, plus the fact that the number of local Romance speakers was much higher, there is far less transfer into Andalusi Arabic from Berber than there is from Romance.

These transfers basically consist of lexical borrowings, which are mainly to be found in Arabic-script botanical glossaries, and have been analysed by various authors, including: Ferrando (1997),[16] Corriente (1981; 1998a; 2002) and Corriente et al. (2017; forthcoming).

[16]This work includes a previously unpublished analysis conducted by G. S. Colin.

3 Contact-induced changes in Andalusi Arabic

3.1 Contact with Andalusi Romance

A special feature of the linguistic history of al-Andalus is that, within a few centuries, a situation of bilingualism, whereby the Romance language was the L1 for most of the population while Andalusi Arabic was L2, was reversed, eventually leading to a third phase of monolingualism using only Andalusi Arabic.

Transfers from Romance to Andalusi Arabic probably took place during the first of the bilingualism phases, a situation which, according to Corriente (2005; 2008), must have lasted two hundred years, from the eighth to the tenth century.

It is difficult to diagnose what type of transfer took place in such an ancient contact situation. When the agents of change used Romance (the source language; SL) as L1 and Andalusi Arabic (the recipient language; RL) as L2, the type of change was imposition, according to the framework of Van Coetsem (1988; 2000). As we have seen, however, this situation would evolve, and the agents of change would come to have Andalusi Arabic (the RL) as their L1 and Romance (the SL) as their L2, meaning that transfer in this situation would be classified as borrowing in Van Coetsem's framework.

However, in cases such as this where the precise sociolinguistic situation at a given time is impossible to judge, it is difficult to establish whether the agents of change had two L1s or one L1 and one L2. Thus, the possibility exists that the contact-induced language changes taking place are a convergence type of transfer (in the terms of Lucas 2015).

3.1.1 Phonology

One contact-induced language change from Romance concerned the prosodic rhythm of Andalusi Arabic. The quantitative rhythm of Old Arabic was replaced by the intense stress system of early Romance languages in the Iberian Peninsula.[17] Thus, while all Old Arabic and Neo-Arabic varieties feature a prosodic rhythm that distinguishes long and short syllables, Andalusi Arabic is the only variety where this quantitative rhythm was replaced by a system where there is no phonemic vowel length (Corriente 1977; 1992b; Corriente et al. 2015: 75–78).

In this case, the agents of change were presumably L1 speakers of Andalusi Romance, making the transfer a case of imposition on the L2, Andalusi Arabic.

[17] A change which had taken place in Latin about one thousand years earlier. This language evolved from a quantitative stress system to an intense stress system in some of its daughter languages. The same process took place later in Andalusi Arabic.

The altered use of the *matres lectionis* in the Arabic script constitutes graphemic evidence of this change in prosodic rhythm. Thus, in Andalusi sources, the graphemes which traditionally mark the Old Arabic long vowels are sometimes used to mark etymologically short vowels, to indicate that these are stressed. For instance: مقاص *muqāṣ* = /muqáṣṣ/ 'pair of scissors' (OA *miqaṣṣ*), أسقوف *usqūf* = /usqúf/ 'bishop' (OA *usquf*), قنفود *qunfūd* = /qunfúd/ 'hedgehog' (OA *qunfud*).

Moreover, historically long vowels that were not stressed are often represented without the regular *matres lectionis*, for instance: فران *firān* = /firán/ 'mice', عم *ʕam* = /ʕam/ 'year'.

Another instance is the very name *al-Andalus*, pronounced by its inhabitants as /alandalús/, a fact known due to the *matres lectionis* for /ū/ which appears in the final syllable, indicating that this syllable is stressed: الاندلوس *al-andalūs* = /alandalús/.

In addition, lexical borrowings from Andalusi Arabic currently found in Ibero-Romance languages also attest to this change of prosodic rhythm. For instance, the Spanish word *andaluz* (stressed on the last syllable) can only originate in the Andalusi word /alandalús/, while the Spanish word *azahar* 'orange blossom' (also stressed on the last syllable) comes from the Andalusi word /azzahár/, rather than directly from Old Arabic *zahr* 'flower'.

The use of *matres lectionis* in this way was by no means systematic, since less cultivated scribes inserted or suppressed them arbitrarily; a fact which could be interpreted as indicative of an incipient evolution towards the loss of the phonological value of stress in Andalusi prosody, (Corriente et al. 2015: 76, fn. 213), a phenomenon that today characterizes Moroccan Arabic, perhaps the last step of this evolution in Maghrebi Arabic dialects.

In some cases, a graphic gemination of the following consonant instead of the grapheme of the vocal quantity is an alternative means of indicating an accentuated vowel, for instance: أسقفف *usquff* = /usqúf/ 'bishop', ثقة *θiqqa* = /θíqa/ 'trust', (Corriente et al. 2015: 77).

Andalusi Arabic also features the appearance of three marginal phonemes /p/, /g/ and /č/ as transferred from Andalusi Romance, which, however, may not have existed in some Andalusi sub-dialects. Bearing in mind that these phonemes were incorporated through loanwords (Corriente 1978), we can assume that the agents of change had Andalusi Arabic as L1 and that therefore this is a borrowing type of transfer. Examples include: *čípp* 'trap', *čiqála* 'cicada', *čírniya* 'blackbird' (Corriente et al. 2015: 57). As these phonemes exist even in late toponymy it may be concluded that they were part of the Andalusi phonological system.

Another contact-induced phonological change was the partial loss of contrastive

velarization in some phonemes. As velarization does not exist in Romance languages, we can assume that this was a case of phonological imposition by L1 Romance speakers on their L2 Andalusi Arabic.

The effects of this change are visible, for instance, in the frequent interchangeability of /s/ and /ṣ/. Recurrent permutations between both realizations exist and pseudo-corrections are also in evidence. For example: /sūr, ṣūr/ 'wall', /nāqūs, nāqūṣ/ 'bell', /qaswa, qaṣwa/ 'cruelty'. This is not, however, a very common feature and took place only in the early stages of the Arabization process (Corriente et al. 2015: 82).

The spirantization of occlusives is another example of contact-induced phonological change in Andalusi Arabic, due imposition from Andalusi Romance. According to Romanists, this phenomenon was commonly found in Romance languages since the Latin period.[18]

The spirantization of /d/ > [ð] can be observed. Authors of Andalusi Arabic would write ⟨ذ⟩ (ð) rather than ⟨د⟩ (d) for both *d and *ð because they considered both sounds to be allophones of /d/, particularly in postvocalic position.[19] The realization of the /d/ phoneme clearly changed through contact with Andalusi Romance. This is a widespread feature noted in various authors, regions, ages and social groups. For instance: جذول /ğaðwal/ 'creek' < *ğadwal*, حفيذ /ḥafīð/ 'nephew' < *ḥafīd*, ألحذ /al-ḥaðð/ 'Sunday' < *al-ḥadd*, سيذي /sīði/ 'my lord' < *sīdi*. This phenomenon seems to have been more common in lower and middle registers of Andalusi.

Another example is the spirantized allophone of /b/, [β], which could constitute a borrowing from Romance or Zenati Berber. This may be confirmed by the use of ⟨f⟩ to represent /b/ (as in قسفورى *qasfūrā* < *kuzbara* 'coriander', فش *fiš* < *baš/biš* 'in order to'), or by confusion between both phonemes: *baysāra*/*faysāra* 'a dish of cooked beans' (Corriente et al. 2015: 19).

3.1.2 Morphology

A noteworthy contact-induced morphological change concerns the elimination of a gender distinction in the second person singular of both pronouns and verbs, as in *taqtúl* 'you kill', *tikassár* 'you break', *taḥtarám* 'you respect', *taḫriğ* 'you

[18]The spirantization of the occlusives is also a feature of some Arabic varieties spoken in Morocco, especially, though not exclusively, in the north (Sánchez & Vicente 2012: 235–236). In this case, the agents of change were Arabic–Berber bilingual speakers who imposed the phonology of their L1 Berber on their L2 Arabic. This may have also happened in Andalusi society, though data to corroborate it is insufficient.

[19]This spirantization is also realized in other positions, however.

throw', see (Corriente et al. 2015: 154–155).

The addition of Romance suffixes to Arabic words to produce hybrid terms was another example of morphological transfer. These suffixes are numerous. For instance, the augmentative suffix *-ūn*, as in *ǧurrún* 'big jar' < *ǧarra* 'jar', *raqadún* 'sleepyhead' < *rāqid* 'asleep', and the agentive suffix *-áyr*, as in *ǧawabáyr* 'cheeky' < *ǧawāb* 'answer' (cf. Corriente 1992a: 126–131; Corriente et al. 2015: 230–231).

3.1.3 Syntax

Changes in gender agreement also arguably result from contact-induced change: *ʕáyn* 'eye', *šáms* 'sun', and *nár* 'fire' are generally feminine in Arabic but were occasionally treated as masculine in Andalusi Arabic, as their translation equivalents are in Romance. Likewise, *má* 'water' and *dwá* 'medicine' are masculine in Arabic but were sometimes considered feminine in Andalusi Arabic, again on a Romance model (Corriente et al. 2015: 232). This was presumably a case of imposition, where the agents of change were L2 speakers of Andalusi Arabic.

There are cases of concordless determination constructions in qualifying syntagms following the Romance construction, for instance: *alʕaqd θānī* 'the second contract' instead of more typical *al-ʕaqd aθ-θānī* (Corriente et al. 2015: 186). These examples come from texts written by bilingual Mozarabs from Toledo; since they were either dominant in Andalusi Arabic or had both Andalusi Arabic and Andalusi Romance as L1s, this change must have been either an instance of borrowing or of convergence.

There are instances of a construction using the analytic genitive with the preposition *min* 'of' as well as innovative uses of *li* 'for'. These are found particularly in late texts with strong influence from Andalusi Romance (cf. Corriente 2012). As in the previous case, we are dealing here with agents of change who are either dominant in Andalusi Arabic and thus borrowing from Andalusi Romance, or this is an instance of convergence brought about by speakers of both languages as L1s.

(2) Late Andalusi Arabic (Corriente et al. 2015: 233–234)

a. mudda min ʕām-ayn
period from year-DU
'a two-year period'

b. min ʕām
from year
'one year old'

c. naḏruğ li wild-ī
go_out.IMPF.1SG to father-OBL.1SG
'I look like my father.'

The examples in (2) are clearly calqued on Romance expressions: *un periodo de dos años, de un año* and *salgo a mi padre*, respectively.

3.1.4 Lexicon

Lexical borrowings from Romance in Andalusi Arabic constitute less than 5%, according to Corriente (1992a: 142).[20]

The most common semantic fields are botanical terms of species endemic to the Iberian Peninsula, as in *ūliyā* 'olive', *amindāl* 'almond', *blāṭur* 'water lily', *bulmuš* 'elm tree', and zoological terms, as in *burrays* 'lamb', *poḥóta* 'whiting', *buṭrah* 'mule', *ṭābaraš* 'capers'. For more examples, see Corriente et al. (2017). Other semantic fields are parts of the body, as in *imlīq* 'navel' and *muğğa* 'breast', family relations, as in *šuqru* 'father-in-law', *šubrīn* 'nephew', and household items and technicalities of various professions, as in *šuqūr* 'axe' and *šayra* 'basket', (Corriente et al. 2015: 224).

Some words even adapted to the pattern of broken plural in Andalusi Arabic, for instance *š(u)nyūr* 'sir', pl. *šanānīr*, though most used the suffix of the regular plural *-āt*.

3.2 Contact with Berber

As with the Arabic–Romance contact situation, lack of information regarding the sociolinguistic status of Berber speakers in al-Andalus in the relevant period makes it difficult to classify the relevant changes according to the types of agentivity involved. That said, since we have no reason to think that significant numbers of native Arabic speakers would have acquired Berber languages as L2s, the changes described here seem most likely to be the result of imposition by L1 Berber speakers.

[20]The number of lexical borrowings from Andalusi Arabic into Romance languages spoken in Spain is larger. According to Corriente (2005), its number is close to 2000, not counting the lexical derivations and place names included by other authors, who have put the number at 4000 or even 5000. Many of the terms in question are nowadays obsolete (Corriente 2005: 203, fn. 59). We must not forget that these languages had a different social status during the period of bilingualism, a major element in contact-induced language changes. In such situations, less prestigious languages always receive a larger number of transfers (cf. Corriente et al. 2019).

3.2.1 Phonology

Available data is always written sources and it is therefore hard to be certain about the existence of contact-induced phonological changes.

The realization of *k as [ḫ] has been considered a Zenati Berber influence (Corriente 1981: 7). For instance: *aḫθar* 'more', *aḫṭubar* 'October' (Corriente et al. 2015: 61).

The replacement of /l/ with /r/, as in Tarifit Berber, could be another instance of transfer from Berber. Thus, the following spellings in documents written in Latin script could be instances of possible assimilation-induced allophones: *Huaraç, Hurad, Uarat* < *walad* 'boy'. The late source where these spellings are found, documents written by Valencian *moriscos* in the second half of the sixteenth century (Labarta 1987), suggests that this change could have been introduced through contact with the last Berber immigration waves into al-Andalus (thirteenth century). However, this trait may not have been generalized in the speech of the wider community, and could merely represent idiolectal variation or even misspelling.

3.2.2 Lexicon

While contact-induced changes in Andalusi Arabic from Berber were initially considered very scarce, more comprehensive analyses of the sources have revealed that changes may not have been so insignificant.[21] In fact, the list compiled by Corriente in 1981 contained 15 Berber loanwords in Andalusi Arabic (1981: 28–29), the list in his dictionary of 1997 listed 62 (Corriente 1997a: 590), and the compilation made by Ferrando the same year included 82, of which 39 corresponded to an unpublished study by G. S. Colin and 43 were compiled from proposals made by various other scholars (Ferrando 1997: 133). The most recent list contains 115 Berber loanwords (Corriente et al. 2017: 1432–1433).

As Ferrando (1997: 140) points out, these borrowings appear mostly in earlier sources, and their number decreases considerably in later sources. This fact could be put down to the social and cultural prestige Andalusi Arabic achieved in later centuries, even contributing to social cohesion and, therefore, linguistic cohesion. Most lexical transfers must have taken place in the early centuries of the existence of al-Andalus, prior to the arrival of new Berber speakers, the Almoravids and the Almohads. For obvious geographical reasons it is quite likely that the Berber-speaking Muslims (already Arabized) who reached the Iberian Peninsula

[21] For instance, linguistic analyses of some sources, such as the botanical glossaries written in al-Andalus, have yielded a large number of Berber loans in Andalusi Arabic (cf. al-Išbīlī 2004; 2007; Corriente 2012).

with the first Muslim troops came from an area in modern northwestern Morocco, the region known as Jbala. Ghomara and Senhaja are the vernacular Berber varieties from this region. These non-Zenati varieties are different from those spoken in the Rif (Kossmann 2017). It is therefore probable that Ghomara and Senhaja Berber were the sources of a good deal of these borrowings, though any attempt at classifying them is hindered by the lack of detailed phonetic or morphological data.

Semantically, most of these lexical borrowings correspond to phytonyms and zoological terms, socio-political symbols and names of weapons, clothing, food, and household goods. The number of Berber loanwords that were regularly used by the Andalusi population is not easily determined, as many are the names of plants probably existing only in Berber botanical treatises.

The following are some examples from Corriente et al. (2017):[22] *azarūd* 'sweet clover' < *azrud*/*aẓrud* , *aṭṭifu* 'take him' < *əṭṭəf* 'take', *āwurmī* 'garden street' < *awurmi*/*iwurmi*, *aɣlāl* 'snails' < *aɣlal*, *tamaɣra* 'banquet' < *tamǝɣra* 'wedding party', *zuɣzal* (with agglutination of the preposition *s-* 'with') 'half-pike (Berber weapon)' < *ugzal*, *tāqra* 'terrine' < *tagra* 'wooden dish to make couscous', *aqrūn* 'pancakes cut into squares and eaten with honey' < *aɣrum* 'bread'.[23]

Some of these loans present a chronological problem. The problematic items are those which have an ungeminated /š/ or /q/, phonemes that were transferred to the Berber varieties through contact with Arabic.[24] These would appear, therefore, to be later loans that arrived with the Berber already Arabized or through Moroccan Arabic, for instance: *išir* < *iššir* 'boy', *finniš* 'mule' < *afənniš* 'snub-nosed',[25] *barqī* < *abərqi* 'slap'.[26]

Some of these loans do not appear in modern dictionaries of Berber varieties,

[22]The Berber origin of some of the lexical borrowings from these lists is only probable, not certain. Due to the characteristics of the sources, written in Arabic or Romance by possibly non-Berber-speaking scribes, the available information sometimes does not allow us to go beyond mere working hypotheses. It is also difficult to decide which Berber variety they belong to: Tarifit, Taqbaylit and Tashelhiyt have all been found. Note also that all Arabic items in this section are rendered as transliterations of their Arabic-script orthography, rather than transcription of their (assumed) phonology.

[23]This item exists in Taqbaylit with the meaning 'unleavened cooked pasta cookie' (Dallet 1982). The ending *-um* becomes *-un* due to a metanalysis that associates it with the Romance suffix *-on*, which is highly productive in Andalusi Romance.

[24]I thank Maarten Kossmann for this and other valuable comments on the section of this work dealing with contact between Andalusi Arabic and Berber varieties.

[25]In Moroccan Arabic *fənnīš*/*fənnūš* (de Prémare 1998: 167).

[26]According to de Prémare (1993: 5), the Moroccan Arabic word *ābāṛəq* 'slap' is also a loanword of Berber origin.

such as *aryīs* 'barberry' < *aryis*,[27] *āðiqal* 'watermelon' < *adigal*, *maqaqūn* 'stallion' < *amaka*.[28]

In some cases we have loans that come from Vulgar Latin to Andalusi Arabic via Berber, for instance: *fullūs* 'chicken' < *afəllus* (Berber) < *pullus* (Vulgar Latin), *bāqya* 'large clay dish' < *tabaqit*/*θabəqqišθ* (Tarifit) 'great dish of superior quality'[29] < *bacchia* (Vulgar Latin) 'goblet, water jug', *hirkāsa* 'rustic leather shoe' < *arkasən* (Kabyle) or *arkas*, *ahərkus* (Tarifit) perhaps < *calcĕus* (Vulgar Latin), *tirfās* 'truffles' < *tərfas* (Berber) < *tuferas* (Vulgar Latin), *zabzīn* 'low-quality couscous' < *zabazin* (Berber, with agglutination of the preposition *s-* 'with') < *pisellum* (Vulgar Latin, diminutive of *pisum* 'pea').[30] These transfers are very likely to have first taken place in North Africa (the northern part of present-day Morocco), since we know that some variety of Vulgar Latin was in contact there with the Berber variety of the region before the arrival of Muslim troops (cf. Heath, this volume). The Berber-speaking Andalusians would have then later transferred these items to Andalusi Arabic.[31]

Some of these lexical borrowings have certain characteristics that demonstrate greater integration than others in Andalusi Arabic:

1. Morphophonemic adaptations.

    a) Phonemic adaptation to Arabic (although this may simply be a problem of orthography, since the Arabic script lacks a means of representing the Berber phonemes /g/ and /ẓ/). /g/ is represented as ⟨k⟩, ⟨q⟩ or ⟨ǧ⟩: *akzal/aqzal* 'pike (characteristic weapon of the Berbers)' < *agzal*;[32] *āðiqal* 'watermelon' < *adigal*; *arǧān* 'argan tree' < *argan*,

[27]The Berber origin of this item has nevertheless been affirmed by Colin and Ferrando, since this circumstance is clear thanks to the data provided by Ibn al-Bayṭār (Ferrando 1997: 110–111). It is documented in Moroccan Arabic, *āryīs* 'barberry' (de Prémare 1995: 151), and in Spanish it has become *alargue* and *alguese*, and in Portuguese *largis* (Corriente et al. forthcoming). A fall into disuse in the SL is perhaps the reason of its absence from the current dictionaries.

[28]The last two lexical borrowings are documented in the Andalusi source *kitābu ʕumdati ṭ-ṭabīb*, by Abu l-Ḫayr al-ʔIšbīlī (2004; 2007), a botanist of the eleventh century. However, their Berber origin is quite doubtful for M. Kossmann (personal communication).

[29]See Ibáñez (1949: 272) whose transcription is *zabeqqixz*.

[30]The word exists in Moroccan Arabic as *ābāzīn* (de Prémare 1993: 5), and in Kabyle Berber as *tabazint* (augmentative of *abazin*).

[31]A number of these Berber loans have then gone on to reach the Romance languages through Andalusi Arabic. The most recent list includes forty of these borrowings in Romance languages (Corriente et al. 2019).

[32]Andalusi Arabic seems to have had a diminutive form of this item: *tagzalt* (modern dictionaries give the diminutive *tagəzzalt* 'small stick'; Taïfi 1991). This could then be the source of Castilian *tragacete* and Portuguese *tragazeite* 'dart' (Corriente et al. forthcoming).

qillīd 'Berber prince' < *agəllid*, while /ẓ/ is represented as ⟨z⟩: *zawzana* 'mutism' < *aẓiẓun*, *lazāz* 'werewolf' < *aẓẓaẓ*.

b) Elimination of typically Berber morphemes: e.g., the loss of prefix *a-* of masculine nouns: *bāzīn* 'a dish of couscous, meat and vegetables' < *abazin*, *dād abyaḍ* 'white chameleon' < *addad*, *mizwār* 'manager, commander' < *amәzwaru* 'first', *finniš* 'mule' < *afәnniš* 'snub-nosed', *mazad* 'Quranic school' < *amzad*. Likewise the loss of prefix and suffix *t-…-t* of feminine nouns: *zaɣnaz* 'brooch, buckle' < *tisәɣnәst* (Tarifit),[33] *muzūra* 'horse braid' < *tamzurt* (Tarifit and Kabyle), *saryant* 'root of the orpine plant' < *tasәryint*, as well as elimination of prefix *t-*, as in *abɣā* 'wild bramble' < *tabɣa* .

2. Another process for the integration of lexical borrowing involves fitting Berber words to Arabic patterns, as in *zawzana* 'mutism' (with the Arabic pattern CawCaCa) < *aẓiẓun*, *harkama* 'tripe stew' (with the Arabic pattern CaCCaCa) < *urkimen*, *hirkāsa* 'rustic leather shoe' (with the Arabic pattern CiCCāCa) < *arkasәn* (Kabyle) or *arkas*, *ahәrkus* (Tarifit).

4 Conclusion

Andalusi Arabic developed in the Iberian Peninsula through intra-Arabic leveling and contact with two other language types: Romance and Berber. This situation spanned a long period of time and resulted in a good deal of contact-induced change.

Initially the L2 of most of the population, after a two-century gestation process, Andalusi Arabic gradually became the dominant language, overtaking the Romance dialect spoken by the local population. The main reason was the growing social prestige attached to Arabic in an Islamic society, in contrast to the lower social status of Andalusi Romance, which first became an L2, before the bilingual situation eventually disappeared. This contact situation resulted in a number of contact-induced changes in all areas of grammar, but it is often difficult to diagnose what type of transfer took place in such an ancient contact situation.

Concerning Berber varieties, modern historiography reveals that the interaction between Arabic-speaking and Berber-speaking populations on both sides of

[33]This is a noun of instrument derived from the verb *ɣnәs* 'to tie with a brooch'. Corriente derives it from *asәgnәs* 'needle', see (Corriente et al. forthcoming), but the phoneme /ɣ/ makes the first option more likely (M. Kossmann, personal communication).

the Strait of Gibraltar facilitated lasting language contact. The role of Berber in the language development of al-Andalus, however, has not yet been analysed in depth. The nature of the available data is such that lexical borrowings are the only transfers that have been well described at present.

Future research would be particularly desirable with regard to contact-induced changes in Andalusi Arabic due to the presence of Berber varieties in the Iberian Peninsula. This should involve collaboration between scholars of Berber and of Arabic.

Further reading

Corriente (1997b) provides a linguistic analysis of Andalusian strophic poetry.
Corriente (2005) offers valuable information concerning the impact of Andalusi Arabic on Ibero-Romance.
Corriente et al. (2015) is the most up-to-date book-length description of Andalusi Arabic grammar. It contains a section dealing with transfer from Romance and Berber.
Ferrando (1997) offers an etymological description of some Berber loanwords in Andalusi Arabic.
Vicente (2010) details the Andalusi influence on the dialects of northern Morocco.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person
DU	dual
IMPF	imperfect (prefix conjugation)
L1	first language
L2	second language
M	masculine
OA	Old Arabic
OBL	oblique
RL	recipient language
SG	singular
SL	source language

References

al-Išbīlī, Abū l-Ḫayr. 2004. *Kitābu ʿumdati ṭṭabīb fī maʿrifati nnabāt likulli labīb: Libro base del médico para el conocimiento de la botánica por todo experto.* Joaquín Bustamante Costa, Federico Corriente & Mohand Tilmatine (eds.). Vol. 1: *Texto árabe.* Madrid: Consejo Superior de Investigaciones Científicas.

al-Išbīlī, Abū l-Ḫayr. 2007. *Kitābu ʿumdati ṭṭabīb fī maʿrifati nnabāt likulli labīb: Libro base del médico para el conocimiento de la botánica por todo experto*. Mohand Tilmatine, Federico Corriente & Joaquín Bustamante Costa (eds.). Vol. 2: *Traducción anotada*. Madrid: Consejo Superior de Investigaciones Científicas.

Al-Wasif, Muhammad-Fajri. 1990. La inmigración de árabes yemeníes a al-Andalus desde la conquista islámica (92/711) hasta fines del siglo II/VIII. *Anaquel de Estudios Árabes* 1. 203–219.

Barceló, Carmen & Ana Labarta (eds.). 2009. *Archivos moriscos: Textos árabes de la minoría islámica valenciana, 1401–1608*. Valencia: Universidad de Valencia.

Chalmeta, Pedro. 2003. *Invasión e islamización: La sumisión de Hispania y la formación de al-Andalus*. Jaén: Universidad de Jaén.

Corriente, Federico. 1977. *A grammatical sketch of the Spanish Arabic dialect bundle*. Madrid: Instituto Hispano-Arabe de Cultura.

Corriente, Federico. 1978. Los fonemas /p/, /č/ y /g/ en árabe hispánico. *Vox Romanica* 37. 214–218.

Corriente, Federico. 1981. Notas de lexicología hispanoárabe (III y IV). III: Los romancismos del Vocabulista. IV: Nuevos berberismos del hispano-árabe. *Awrāq* 4. 5–30.

Corriente, Federico. 1992a. *Arabe andalusí y lenguas romances*. Madrid: Editorial MAPFRE.

Corriente, Federico. 1992b. Linguistic interferences between Arabic and the Romance languages of the Iberian Peninsula. In Salma Khadra Jayyusi (ed.), *The legacy of Muslim Spain*, 443–456. Leiden: Brill.

Corriente, Federico. 1995. El idiolecto romance andalusí reflejado por las *xarajāt*. *Revista de Filología Española* 75. 5–33.

Corriente, Federico. 1997a. *A dictionary of Andalusi Arabic*. Leiden: Brill.

Corriente, Federico. 1997b. *Poesía dialectal árabe y romance en Alandalus: Cejeles y* xarajāt *de* muwaššaḥāt. Madrid: Gredos.

Corriente, Federico. 1998a. Le berbère à Al-Andalus. *Études et Documents Berbères* 15–16. 269–275.

Corriente, Federico. 1998b. On some features of late Granadian Arabic (mostly stress). In Jordi Aguadé, Patrice Cressier & Ángeles Vicente (eds.), *Peuplement et arabisation au Maghreb occidental: Dialectologie et histoire*, 53–57. Madrid, Zaragoza: Casa de Velázquez, Universidad de Zaragoza.

Corriente, Federico. 2000. El romandalusí reflejado por el glosario botánico de Abulxayr. *Estudios de Dialectología Norteafricana y Andalusí* 5. 93–241.

Corriente, Federico. 2002. The Berber adstratum of Andalusi Arabic. 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*, 105–111. Wiesbaden: Harrassowitz.

Corriente, Federico. 2005. El elemento árabe en la historia lingüística peninsular: Actuación directa e indirecta. Los arabismos en los romances peninsulares (especialmente en castellano). In Rafael Cano Aguilar (ed.), *Historia de la lengua española*, 185–206. Barcelona: Ariel.

Corriente, Federico. 2008. *Romania arábica: Tres cuestiones básicas: Arabismos, "mozárabe" y "jarchas"*. Madrid: Editorial Trotta.

Corriente, Federico. 2012. Une première approche des mots occidentaux (arabes, berbères et romans) dans l'ouvrage botanique d'As-Sarif Al-Idrisi. 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 elèves et collègues*, 57–63. Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo (IEIOP).

Corriente, Federico, Christophe Pereira & Ángeles Vicente. 2015. *Aperçu grammatical du faisceau dialectal arabe andalou: Perspectives synchroniques, diachroniques et panchroniques* (Encyclopédie Linguistique d'Al-Andalus, vol. 1). Berlin: De Gruyter.

Corriente, Federico, Christophe Pereira & Ángeles Vicente. 2017. *Dictionnaire du faisceau dialectal arabe andalou: Perspectives phraséologiques et étymologiques* (Encyclopédie Linguistique d'Al-Andalus, vol. 2). Berlin: De Gruyter.

Corriente, Federico, Christophe Pereira & Ángeles Vicente. 2019. *Dictionnaire des emprunts ibéro-romans: Emprunts à l'arabe et aux langues du Monde Islamique* (Encyclopédie Linguistique d'Al-Andalus, vol. 3). Berlin: De Gruyter.

Corriente, Federico, Christophe Pereira & Ángeles Vicente. Forthcoming. *Le substrat roman et l'adstrat berbère du faisceau dialectal andalou* (Encyclopédie Linguistique d'Al-Andalus, vol. 4). Berlin: De Gruyter.

Dallet, Jean-Marie. 1982. *Dictionnaire kabyle-français, français-kabyle*. Paris: SELAF.

Ferrando, Ignacio. 1995. Los romancismos de los documentos mozárabes de Toledo. *Anaquel de Estudios Árabes* 6. 71–86.

Ferrando, Ignacio. 1997. G. S. Colin y los berberismos del árabe andalusí. *Estudios de Dialectología Norteafricana y Andalusí* 2. 105–146.

Fierro Bello, María Isabel. 2001. *Al-Andalus: Saberes e intercambios culturales*. Barcelona: Icaria.

Galmés de Fuentes, Álvaro. 1994. *Las jarchas mozárabes*. Barcelona: Crítica.

Guichard, Pierre. 1989. Otra vez sobre un viejo problema: Orientalismo y occidentalismo en la civilización de la España Musulmana. In Generalitat Valenciana (ed.), *En torno al 750 aniversario: Antecedentes y consecuencias de la conquista de Valencia*, 73–96. Valencia: Generalitat Valenciana.

Guichard, Pierre. 1995. *Al-Andalus: Estructura antropológica de una sociedad islámica en Occidente*. Granada: Servicio de Publicaciones de la Universidad. Facsimile edition prepared by Antonio Malpica Cuello. Originally published in 1976 by Barral.

Ibáñez, Esteban. 1949. *Diccionario rifeño-español (etimológico)*. Madrid: Instituto de Estudios Africanos.

Kossmann, Maarten. 2017. La place du parler des Senhaja de Sraïr dans la dialectologie berbère. In Ángeles Vicente, Dominique Caubet & Amina Naciri-Azzouz (eds.), *La région du Nord-Ouest marocain: Parlers et pratiques sociales et culturelles*. Zaragoza: Prensas de la Universidad de Zaragoza.

Labarta, Ana. 1987. *La onomástica de los moriscos valencianos*. Madrid: Consejo Superior de Investigaciones Cientificas.

Lentin, Jérôme. 2011. Middle Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Lucas, Christopher. 2015. Contact-induced language change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.

Manzano Moreno, Eduardo. 1990. Beréberes de al-Andalus: Los factores de una evolución histórica. *Al-Qantara* 11(2). 397–428.

Menéndez Pidal, Ramón & Alvaro Galmés de Fuentes. 2001. *Islam y cristiandad: España entre las dos culturas*. Málaga: Universidad de Málaga.

de Prémare, André-Louis. 1993. *Dictionnaire arabe-français*. Vol. 1. Paris: l'Harmattan.

de Prémare, André-Louis. 1995. *Dictionnaire arabe-français*. Vol. 5. Paris: l'Harmattan.

de Prémare, André-Louis. 1998. *Dictionnaire arabe-français*. Vol. 10. Paris: l'Harmattan.

Sánchez, Pablo & Ángeles Vicente. 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 elèves et collègues*, 223–252. Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo (IEIOP).

Taïfi, Miloud. 1991. *Dictionnaire tamazight-français (Parlers du Maroc central)*. Paris: L'Harmattan.

Terés Sádaba, Elías. 1957. Linajes árabes en Al-Andalus según la "Yamhara" de Ibn Ḥazm. *Al-Andalus* 22. 55–111.

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. 2006. *El proceso de arabización de Alandalús: Un caso medieval de interaccíon de lenguas*. Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo (IEIOP).

Vicente, Ángeles. 2010. Andalusi influence on northern Morocco following various centuries of linguistic interference. In Juan Pedro Monferrer-Sala & Nader Al Jallad (eds.), *The Arabic language across the ages*, 141–159. Wiesbaden: Reichert.

Wasserstein, David. 1991. The language situation in al-Andalus. In *Studies on the* Muwaššaḥ *and the* Kharja*: Proceedings of the Exeter international colloquium*, 1–15. Oxford: Ithaca Press.

Chapter 5

Ḥassāniyya Arabic

Catherine Taine-Cheikh
CNRS, LACITO

The area where Ḥassāniyya is spoken, located on the outskirts of the Arab world, is contiguous with those of several languages that do not belong to the Afro-Asiatic phylum. However, the greatest influence on the evolution of Ḥassāniyya has been its contact with Berber and Classical Arabic. Loanwords from those languages are distinguished by specific features that have enriched and developed the phonological and morphological system of Ḥassāniyya. In other respects, Ḥassāniyya and Zenaga are currently in a state of either parallel evolution or reciprocal exchanges.

1 Current state and historical development

1.1 Historical development of Ḥassāniyya

The arrival in Morocco of the Banī Maʕqil, travelling companions of the Banī Hilāl and Banī Sulaym, is dated to the thirteenth century. However, the gradual shift to the territories further south of one of their branches – that of the Banī Ḥassān, the origin of the name given to the dialect described here – began closer to the start of the subsequent century.

At that time, the Sahel region of West Africa was inhabited by different communities: on the one hand there were the "white" nomadic Berber-speaking tribes, on the other hand, the sedentary "black" communities.

During the following centuries, particularly during the seventeenth and eighteenth centuries, the sphere of Zenaga Berber gradually diminished, until it ceased to exist in the 1950s, other than in a few tribes in the southwest of Mauritania. At the same time, Ḥassāniyya Arabic became the language of the nomads of the west Saharan group, maintaining a remarkable unity (Taine-Cheikh 2016; 2018b).

Catherine Taine-Cheikh. 2020. Ḥassāniyya Arabic. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 103–123. Berlin: Language Science Press. DOI:??

There is virtually no direct documentation of the region's linguistic history during these centuries. This absence of information itself suggests a very gradual transformation and an extended period of bilingualism.

Despite the lack of documentation of the transfer phenomenon, it seems highly likely that bilinguals played a very important role in the changes described in this chapter.

1.2 Current situation of Ḥassāniyya

The presence of significant Ḥassāniyya-speaking communities is recognized in six countries. With the exception of Senegal and especially of Niger, the regions occupied by these communities, more or less adjacent, are situated primarily in Mauritania, in the north, northeast and east of the country.

The greatest number of Ḥassāniyya-speakers (approximately 2.8 out of a total of four million) are found in Mauritania, where they constitute the majority of the population (approximately 75%). The Ḥassāniyya language tends to fulfil the role of the lingua franca without, however, having genuine official recognition beyond, or even equal to, that which it has acquired (often recently) in neighbouring countries.

2 Contact languages

2.1 Contact with other Arabic varieties

The Islamization of the Ḥassāniyya-speaking population took place at an early date, and Ḥassāniyya has therefore had lengthy exposure to Classical Arabic. For many centuries this contact remained superficial, however, except for among the Marabout tribes, where proficiency in literary Arabic was quite widespread and in some cases almost total. The teaching of Islamic sciences in other places reached quite exceptional levels in certain *mḥāḍ̊ər* (a type of traditional desert university).[1] In the post-colonial era, the choice of Arabic as official language, and the widespread Arabization of education, media and services, greatly increased the Ḥassāniyya-speaking population's contact with literary Arabic (including in its Modern Standard form), though perfect fluency was not achieved, even among the young and educated populations.

[1]These may be referred to as universities both in terms of the standard of teaching and the length of students' studies. They were, however, small-scale, local affairs, located either in nomadic encampments or in ancient caravan cities.

Excluding the limited influence of the Egyptian and Lebanese–Syrian dialects used by the media, the Arabic dialects with which Ḥassāniyya comes in to contact with most often today are those of the neighbouring countries (southern Moroccan and southern Algerian). Most recently Moroccan koiné Arabic has established a presence in the Western Sahara, since the region came under the administration of Morocco.

2.2 Contact with Berber languages

Ḥassāniyya has always been in contact with Berber languages. Currently, speakers of Ḥassāniyya are primarily in contact with Tašlḥiyt (south Morocco), Tuareg (Malian Sahara and the Timbuktu region) and Zenaga (southwest Mauritania). In these areas, some speakers are bilingual in Ḥassāniyya and Berber.

In Mauritania, where Zenaga previously occupied a much larger area, Berber clearly appears as a substrate.

2.3 Contact with languages of the Sahel

Contacts between Ḥassāniyya speakers and the languages spoken in the Sahel have varied across regions and over time, but have left few clearly discernible traces on Ḥassāniyya.

The contact with Soninke is ancient (cf. the toponym *Chinguetti* < Soninke *sí-n-gèdé* 'horse well'), but the effects are hardly noticeable outside of the old cities of Mauritania. The contact with Songhay is both very old and still ongoing, but is limited to the eastern part of the region in which Ḥassāniyya is spoken (especially the region of Timbuktu).

The influence of Wolof, albeit marginal, has always been more substantial in southwestern Mauritania, especially among the Awlād Ban^{y}ūg of the Rosso region. It peaked in the years 1950–70, in connection with the immigration to Senegal of many Moors (e.g. *gord^{y}igen* 'homosexual', lit. 'man-woman'). In Mauritania, the influence of Wolof can still be heard in some areas of urban crafts (e.g. mechanics, electricity), but it is primarily a vehicle for borrowing from French.

Although Pulaar speakers constitute the second-largest linguistic community of Mauritania, contact between Ḥassāniyya and Pulaar is very limited, with the exception of some bilingual groups (especially among the Harratins) in the Senegal River valley.

Certain communities (particularly among the Fulani) were traditionally known for their perfect mastery of Ḥassāniyya. As a result of migration into major cities and the aggressive Arabization policy led by the authorities, Ḥassāniyya has

gained ground among all the non-Arabic speakers of Mauritania (especially in the big cities and among younger people), but this has come at the cost of a sometimes very negative attitude towards the language.

2.4 Contact with Indo-European Languages

Exposure to French has prevailed in all the countries of the region, the only exception being the Western Sahara, which, from the end of the nineteenth century until 1975, was under Spanish occupation.

In Mauritania the French occupation came relatively late and was relatively insignificant. However, the influence of the colonizers' language continued well after the country proclaimed its independence in 1960. That said, it has tended to regress since the end of the twentieth century (especially with the rise of Standard Arabic, e.g. *minəstr* has been replaced by *wazīr* 'minister'), whilst exposure to English has become somewhat more significant, at least in the better educated sections of the population.

3 Contact-induced changes in Ḥassāniyya

3.1 Phonology

3.1.1 Consonants

3.1.1.1 The consonant /ḍ/

As in other Bedouin dialects, /ð̣/ is the normal equivalent of the ⟨ض⟩ of Classical Arabic (e.g. *ð̣mər* 'to have an empty stomach' (CA *ḍamira*) and *ð̣ḥak* 'to laugh (CA *ḍaḥika*). Nonetheless, /ḍ/ is found in a number of lexemes in Ḥassāniyya.

The form [ḍ] sometimes occurs as a phonetic realization of /d/ simply due to contact with an emphatic consonant (compare *ṣḍam* 'to upset' and *ṣadma* 'annoyance', CA *√ṣdm*). However, /ḍ/ generally appears in the lexemes borrowed from Standard Arabic, either in all words of a root, or in a subset of them, for example: *staḥḍaṛ* 'to be in agony' and *ḥaḍari* 'urbanite' but *ḥð̣aṛ* 'to be present' and *maḥʾð̣ṛa* 'Quranic school'. The opposition /ḍ/ vs. /ð̣/ can therefore distinguish a classical meaning from a dialectal meaning: compare *staḥḍaṛ* to *staḥð̣aṛ* 'to remember'.

/ḍ/ is common in the vocabulary of the literate. The less educated speakers sometimes replace /ḍ/ with /ð̣/ (as in *qāð̣i* for *qāḍi* 'judge'), but the stop realization is stable in many lexemes, including in loanwords not related to religion, such as *ḍʕīv* 'weak'.

The presence of the same phoneme /ḍ/ in Berber might have facilitated the preservation of its counterpart in Standard Arabic loans, even though in Zenaga /ḍ/ is often fricative (intervocalically). Moreover, the /ḍ/ of Berber is normally devoiced in word-final position in Ḥassāniyya, just as in other Maghrebi dialects, for example: *ṣayvaṭ* 'to say goodbye', from Berber *√fḍ* 'to send'.

3.1.1.2 The consonant /ẓ/

/ẓ/ is one of the two emphatic phonemes of proto-Berber. This emphatic sibilant sound regularly passes from the source language to the recipient language when Berber words are used in Ḥassāniyya. For example: *aẓẓ* ~ *āẓẓ* 'wild pearl millet' (Zenaga *īẓi*).

However, /ẓ/ is also present in lexemes of a different origin. Among Ḥassāniyya roots also attested in Classical Arabic, *z often becomes /ẓ/ in the environment of /ṛ/ (e.g. *ṛāẓ* 'to try', CA *rāza*; *ṛaẓẓa* 'lightning', CA *rizz*; *ẓəbṛa* 'anvil', CA *zubra*). Sometimes /ẓ/ appears in lexemes with a pejorative connotation, e.g. *ẓṛaṭ* 'fart; lie' (CA *ḍaraṭa*), *ẓagg* 'make droppings (birds)' (CA *zaqq*).

3.1.1.3 The consonant /q/

The normal equivalent to the ⟨ق⟩ of Classical Arabic is the velar stop /g/, as in other Bedouin dialects (e.g. *bagṛa* 'cow', CA *baqara*). However, /q/ is in no way rare.

First of all, /q/ appears, like /ḍ/, in a number of words borrowed from Classical Arabic by the literate: *ʕaqʾd* 'religious marriage contract'; *vassaq* 'to pervert'. The opposition /g/ vs. /q/ can therefore produce two families of words, such as *qibla* 'Qibla, direction of Mecca' and *gəbla* 'one of the cardinal directions (south, southwest or west, depending on the region)'. It can also create a distinction between the concrete meaning (with /g/) and the abstract meaning (with /q/): *θgāl* 'become heavy', *θqāl* 'become painful'.

Next, /q/ is present in several lexemes of non-Arabic origin, such as *bsaq* 'silo', *mzawṛaq* 'very diluted (of tea)', (in southwest Mauritania) *səṛqəlla* 'Soninke people', (in Néma) *sasundaqa* 'circumcision ceremony', (in Walata) *raqansak* 'decorative pattern', *asanqās* 'pipe plunger', *sayqad* 'shouting in public', and (in the southeast) *šayqa* 'to move sideways'. These lexemes, often rare and very local in use, seem to be borrowed mostly from the languages of the Sahel region.[2]

[2] I am currently unable to specify the origin of these terms except that *bsaq* (attested in Zenaga) could be of Wolof origin.

Finally, /q/ is the outcome of *ɣ in cases of gemination, (/ɣɣ/ > [qq]): compare *raqqad* 'to make porridge' to *rɣīda* 'a variety of porridge' (CA *raɣīda*). This correlation, attested in Zenaga and more generally in Berber, can be attributed to the substrate.

Insofar as the contrast between /ɣ/ and /q/ is poorly established in Berber, the substrate could also explain the tendency, sometimes observed in the southwest, to velarize non-classical instances of /q/ (or at least instances not identified as classical): hence *ɣandīr* 'candle' for *qandīr* < CA *qandīl* – this is despite the fact that the shift /ɣ/ > /ʔ/ is very common in Zenaga. However, the influence of Berber does not explain the systematic shift of /ɣ/ to /q/ throughout the eastern part of the Ḥassāniyya region (including Mali): thus eastern *qlab* 'defeat' for southwestern *ɣlab* (CA *ɣalaba*).[3]

3.1.1.4 Glottal stop

The glottal stop is one of the phonemes of Zenaga (its presence in the language is in fact a feature that is unique among Berber varieties), however it is not found in Ḥassāniyya, with the exception of words borrowed from Standard Arabic, e.g. *tʔabbad* 'to live religiously', *danāʔa* 'baseness' and *taʔḥīr* 'postponement'. Very rarely the glottal stop is also maintained when it occurs at the end of a word as in *baṛṛaʔ* 'to declare innocent'.

3.1.1.5 Palatalized consonants

There are three palatalized consonants: two dental (/tʸ/ and /dʸ/) and a nasal /nʸ/. Unlike the phonemes discussed above, these are very rare in Ḥassāniyya, especially /nʸ/.

The palatalized consonants are also attested in certain neighbouring languages of the Sahel, as well as in Zenaga (but these are not phonemes of Common Berber). They are rather infrequent in the Zenaga lexicon, occurring especially in syntagmatic contexts (*-d+y-*, *-n+y-*) and in morphological derivation (formation of the passive by affixation of a geminate /tʸ/).

In Ḥassāniyya, the palatalized consonants mostly appear in words borrowed from Zenaga or languages of the Sahel. Interestingly, certain loanwords from Zenaga are ultimately of Arabic origin and constitute examples of phonological integration, as in *tʸfāɣa*, a given name and, in the plural, the name of a tribe

[3]The regular passage from /ɣ/ to /q/ is a typical Bedouin trait, related to the voiced realization (/g/) of *q. It occurs especially in southern Algeria, in various dialects of the Chad–Sudanese area, and in some Eastern dialects (Cantineau 1960: 72).

< Zenaga *atʸfāɣa* 'marabout' < CA *al-faqīh*, and *ḫurūdʸ* 'leave (from Quranic school)' < Zenaga *ḫurūdʸ* < CA *ḫuruǧ* 'exit'.

One should also note the palatization of /t/ in certain lexemes from particular semantic domains (such as the two verbs related to fighting *tʸbəl* 'to hit hard' and *kawtʸam* 'boxer'). This may suggest the choice of a palatalized consonant for its expressive value (and would then be a marginal case of phonosymbolism).

3.1.1.6 Labial and labiovelar consonants

The labiovelar consonants /mʷ, bʷ, fʷ, vʷ/ or /ṃ, ḅ, f̣, ṿ/) are common in Ḥassāniyya, as they are in Zenaga. In both cases, they often come in tandem with a realization [u] of the phoneme /ə/.

This phenomenon may have originally arisen in Zenaga, since the Ḥassāniyya of Mali (where it was most likely in contact with other languages) exhibits greater preservation of a [u] vowel sound and, at the same time, less pronounced labiovelarization of consonants.

The Ḥassāniyya of Mali also has a voiceless use of the phoneme /f/, where the Ḥassāniyya of Mauritania is characterized by the use of /v/ in its place (Heath 2004; an observation that my own studies have confirmed). This phonetic trait does not come directly from Zenaga (in which /v/ exists but is very rare). However, it could be connected with the preference for voiced phonemes in Berber generally and in Zenaga in particular.

3.1.2 Syllabic structures

In Ḥassāniyya, Arabic-derived syllabic structures do not contain short vowels in word-internal open syllables, with the exception of particular cases such as passive participles in *mu-* (*mudagdag* 'broken') and certain nouns of action (*ḥašy* > *ḥaši* 'filling'). However, loanwords from literary Arabic and other languages (notably Berber and French) display short vowels quite systematically in this context: *abadan* 'never' and *ḥazīn* 'sad' (from Standard Arabic); *tamāt* 'gum' (from Zenaga *taʔmað*); *taṃāta* 'tomato'. In fact, it may be noted that, unlike the majority of Berber varieties (particularly in the north), Zenaga has a relatively substantial number of lexical items with short vowels (including *ə*) in open syllables: *kaṛaḍ* 'three', *tuðuṃaʔn* 'a few drops of rain' *awayan* 'languages', *əgəðih* 'necklace made from plants'.[4]

[4]It is precisely for this reason that, regarding the loss of the short vowels in open syllables, I deem the hypothesis of a parallel evolution of syllabic structures in Maghrebi Arabic and Berber to be more convincing than the frequently held alternative hypothesis of a one-way influence of

Furthermore, a long vowel *ā* occurs word-finally in loaned nouns which in Standard Arabic end with *-ā?*: *vidā/vidāy* 'ransom'. In other cases, underlyingly long word-final vowels are only pronounced long when non-final in a genitive construct.

3.2 Morphology

3.2.1 Nominal morphology

3.2.1.1 Standard forms

Nouns and adjectives borrowed from Standard Arabic may often be identified by the presence of: a) open syllables with short vowels, e.g. *vaḍalāt* 'rest of a meal', *ɣaḍab* 'anger', *vasād* 'alteration', *ḥtimāl* 'possibility', b) short vowels /i/ (less frequently /u/) in a closed syllable: *miḥrāb* 'mihrab', *muḥarrir* 'inspector; editor'.

Some syllables are only attested in loanwords, such as the the nominal pattern CVCC, where the pronunciation of the double coda necessitates the insertion of a supporting vowel, in which case the dialect takes on the form CCVC: compare *ʕaqᵊd* 'religious marriage' with *ʕqal* 'wisdom'.

The most characteristic loanword pattern, however, is that of *taḥrīr* 'liberation; verification (of an account)'. In Ḥassāniyya the equivalent of the pattern taCCīC is təCCāC. For the root *√ḥrr*, this provides a verbal noun for other meanings of the verb *ḥarṛaṛ: təḥṛāṛ* 'whipping of wool (to untangle it); adding flour to make dumplings'. As for the form taCaCCuC, the /u/ is sometimes lengthened: *taḥammul* 'obligation', but *tavakkūr* 'contemplation'.

3.2.1.2 Berber affixes

Nouns borrowed from Berber are characterized by the frequent presence of the vowels /a, ā, i, ī, u, ū/. These are of varying lengths, except that in a word-final closed syllable they are always long and stressed. Since these vowels appear in all types of syllables – open and closed – this results in much more varied syllabic patterns than in nouns of Arabic origin.

These loans are also characterized by the presence of affixes which, in the source language, are markers of gender and/or number: the prefix *a/ā-* or *i/ī-* for the masculine, to which the prefix *t-* is also added for the feminine or, more frequently (especially in the singular), a circumfix *t-...-t*. Compare *iggīw* ~ *īggīw*

the Berber substrate on the Arabic adstrate.

'griot' with the feminine form *tiggiwīt ~ tīggīwīt*. A suffix in *-(ə)n* characterizes the plurals of these loanwords which, moreover, differ from the singulars in terms of their vocalic form: *iggāwən ~ īggāwən* 'griots', feminine *tiggawātən ~ tīggawātən*. The presence of these affixes generally precludes the presence of the definite article.

Though these affixes pass from the source language to the target language along with the stems, the syllabic and vocalic patterns of such loans are often particular to Ḥassāniyya: compare Ḥassāniyya *āršān*, plural *īršyūn ~ īršīwən* 'shallow pit' with Zenaga *aʔraš*, plural *aʔraššan* (see Taine-Cheikh 1997a).

Ḥassāniyya speakers whose mother tongue is Zenaga have most likely played a role in the transfer of these affixes and their affixation to nouns of all origins (including those of Arabic origin: a possible example being *tasūvra* 'large decorated leather bag for travelling', cf. *sāvər* 'to travel'). The forms that these speakers use can also be different from those used by other Ḥassāniyya speakers – especially if the latter have not been in contact with Berber speakers for a long time.

It is not proven that Berber speakers are the only ones to have created and imposed these forms which are more Berberized than authentically Berber. However, it may be noted that the gender of nouns borrowed from Berber is generally well preserved in Ḥassāniyya, even for the feminine nouns losing their final *-t*, other than in special cases such as the collective *tayšəṭ* 'thorny tree (*Balanites aegyptiaca*)' with a final *-ṭ* (< Zenaga *tayšaḌ* for *tayšaḍt*).[5] In fact, this indicates a deep penetration of the meaning of these affixes and of Berber morphology in general (up to and including the incompatibility of these affixes with the definite article).

The borrowing of the formants *ən-* 'he of' and *tən-* 'she of' (quasi-equivalents of the Arabic-derived *bū-* and *ūṃ(ṃ)-*) is fairly widespread, in particular in the formation of proper nouns. It is also mostly in toponyms and anthroponyms that the diminutive form with prefix *aɣ-* and suffix *-t* is found, e.g. the toponym *Agjoujt* (< *aɣ-žoʔž-t* 'small ditch').

3.2.2 Verbal Morphology

3.2.2.1 The derivation of *sa-*

The existence of verb forms with the prefix *sa-* is one of the unique characteristics of Ḥassāniyya (Cohen 1963; Taine-Cheikh 2003). There is nothing, however,

[5] In Zenaga, non-intervocalic geminates are distinguished not by length, but rather by tension, and it is this that is indicated by the use of uppercase for the final *Ḍ*.

to indicate that the prefix is an ancient Semitic feature that Ḥassāniyya has preserved since its earliest days. Instead, the regular correspondences between the three series of derived verb forms (causative–factitive vs. reflexive vs. passive) and the specialization of the morpheme *t* as a specific marker of reflexivity underlie the creation of causative–factitives with *sa-*. Neologisms with *sa-* generally appear when forms with the prefix *sta-* have a particular meaning: *staslaʕ* 'to get worse (an injury)' – *saslaʕ* 'to worsen (injury)'; *stabṛak* 'to seek blessings' – *sabṛak* 'to give a blessing'; *stagwa* 'to behave as a griot' – *sagwa* 'to make someone a griot'; *staqbal* 'to head towards the Qibla' – *saqbal* 'to turn an animal for slaughter in the direction of the Qibla'.

Furthermore, the influence of Berber has certainly played a role, since the prefix *s(a)-* (or one of its variants) very regularly forms the causative–factitive structure in this branch of the Afro-Asiatic language family.

In Zenaga, the most frequent realization of this prefix is with a palato-alveolar shibilant, but a sibilant realization also occurs, particularly with roots of Arabic origin. For example: Hass. *sādəb* (variant of *ddəb*) – Zen. *yassiʔðab* 'to train an animal (with a saddle)' < CA *√ʔdb* (cf. *ʔaddaba* 'educate, carefully bring up'); Hass. *sasla* – Zen. *yassaslah* 'to let a hide soak to give it a consistency similar to a placenta' and Hass. *stasla* – Zen. *staslah* 'start to lose fur (of hides left to soak)' < CA *√sly* (cf. *salā* 'placenta').

Parallel to these examples where the Berber forms (at least those with the prefix *st(a)-*) are most likely themselves borrowed, we also find patterns with *sa-/ša-* which are incontestably of Berber origin: compare Ḥassāniyya *niyyər* 'to have a good sense of direction', *sanyar* 'to show the way', *stanyar* 'to know well how to orient oneself' and Tuareg *ener* 'to guide', *sener* 'to make guide'. Typically, however, when Ḥassāniyya borrows causative forms from Berber, it usually integrates the Berber prefix as part of the Ḥassāniyya root, making it the first radical of a quadriliteral root, e.g. Hass. *sadba* – Tuareg *sidou* 'to make s.o. leave in the afternoon' and Hass. *ssadba (< tsadba)* – Tuareg *adou* 'to leave in the afternoon'.

The parallelism between Arabic and Berber is not necessarily respected in all cases, but the forms with initial *s-/š-* are usually causative or factitive in both cases. The only exception concerns certain Zenaga verbal forms which have become irregular upon contact with Ḥassāniyya: thus *yassəðbah* 'to leave in the afternoon' or *yišnar* 'to orient oneself' (a variant of *yinar*), of which the original causative value is now carried by a form with a double prefix (*ž+š*): *yažəšnar* 'to guide'.

3.2.2.2 The Derivation of *u-*

The existence of a passive verbal prefix *u-* for quadrilateral verbs and derived forms constitutes another unique feature of Ḥassāniyya. For example: *udagdag*, passive of *dagdag* 'to break'; *uṭabbab*, passive of *ṭabbab* 'to train (an animal)'; *udāya*, passive of *dāya* 'to cheat (in a game)'.

The development of passives with *u-* was most likely influenced by Classical Arabic, since here the passives of all verbal measures feature /u/ in the first syllable in both the perfect and the imperfect, e.g. *fuʕila*, *yufʕalu*; *fuʕʕila*, *yufaʕʕalu*; and *fūʕila*, *yufāʕalu*, the respective passives of *faʕala*, *faʕʕala* and *fāʕala*.

However, influence from Berber cannot be excluded here since, in Zenaga, the formation of passives with the prefix T^{y} is directly parallel to those of the passives with *u-* in Ḥassāniyya. Moreover, this prefix is *t(t)u-* or *t(t)w-* in other Berber varieties (especially those of Morocco) and this could also have had an influence on the emergence of the prefix *u-*.

3.3 Syntax

3.3.1 Ḥassāniyya–Zenaga parallelisms

Ḥassāniyya and Zenaga have numerous common features, and this is especially true in the realm of syntax. In general, the reason for these common traits is that they both belong to the Afro-Asiatic family and remain conservative in various respects; for example, in their lack of a discontinuous negative construction.

There are, however, also features of several varieties of both languages documented in Mauritania that represent parallel innovations. Thus, corresponding to the diminutive forms particular to Zenaga, we have in Ḥassāniyya *mutatis mutandis* a remarkably similar extension to verbs of the diminutive pattern with infix *-ay-*, e.g. *mayllas*, diminutive of *mallas* 'to smooth over' (Taine-Cheikh 2008a: 123–124).

In the case of aspectual–temporal forms, there are frequent parallels, such as Ḥassāniyya *mā tla* and Zenaga *war yiššiy* 'no longer', Ḥassāniyya *ma-zāl* and Zenaga *yaššiy* 'still', Ḥassāniyya *tamm* and Zenaga *yuktay* 'to continue to', Ḥassāniyya *ʕgab* and Zenaga *yaggara* 'to end up doing'. One of the most notable parallel innovations, however, concerns the future morpheme: Ḥassāniyya *lāhi* (invariable participle of an otherwise obsolete verb, but compare *ltha* 'to pass one's time') and Zenaga *yanhāya* (a conjugated verb also meaning 'to busy oneself with something', in addition to its future function). In both cases we have forms related to Classical Arabic *lahā* 'to amuse oneself', with the Zenaga form

apparently being a borrowing. It seems, therefore, that this borrowing preceded the *lāhi* of Ḥassāniyya and likely then influenced its adoption as a future tense marker. Note also that in the Arabic dialect of the Jews of Algiers, *lāti* is a durative present tense marker (Cohen 1924: 221; Taine-Cheikh 2004: 224; Taine-Cheikh 2008a: 126–127; Taine-Cheikh 2009: 99).

Ḥassāniyya and Zenaga also display common features with regard to complex phrases. For example, concerning completives, Zenaga differs from other Berber languages in its highly developed usage of *ad ~ að*, and in particular in the grammaticalized usage of this demonstrative as a quotative particle after verbs of speaking and thinking (Taine-Cheikh 2010b). This may have had an influence on the usage of the conjunctions *an(n)-* and *ʕan-* (the two forms tend to be confused) in the same function in Ḥassāniyya.

Finally, regarding the variable appearance of a resumptive pronoun in Ḥassāniyya object relative clauses, if influence from Berber (where a resumptive pronoun is always absent) has played any role here, it has simply been to reinforce a construction already attested in the earliest Arabic, whereby the resumptive pronoun is absent if the antecedent is definite, as in (1).

(1) nṛədd ʕlī-kum əṛ-ṛwāye lli ṛadd-Ø ʕlī-ya
tell.IMPF.1SG on-2PL DEF-story REL tell.PRF.3SG.M-Ø on-OBL.1SG
muḥammad
Mohammed
'I am going to tell you the story that Mohammed told me.'

3.3.2 Regional influence of Maghrebi Arabic

The Ḥassāniyya spoken in the south of Morocco is rather heavily influenced by other Arabic varieties spoken in the region. Even among those who conserve virtually all the characteristic features of Ḥassāniyya (preservation of interdentals, synthetic genitive construction, absence of the pre-verbal particle *kā-* or *tā-*, absence of discontinuous negation, absence of the indefinite article), particular features of the Moroccan Arabic koiné appear either occasionally or regularly among certain speakers. The most common such features are perhaps the genitive particle *dyal* (Taine-Cheikh 1997b: 98) and the preverbal particle *kā* (Aguadé 1998: 211, §37; 213, §42).

In the Ḥassāniyya of Mali, usage of a genitive particle remains marginal, although Heath (2004: 162) highlights a few uses of genitive *(n)tāʕ* in his texts.

3.4 Lexicon

3.4.1 Confirmed loanwords

3.4.1.1 Loanwords from Standard Arabic

Verbs loaned from Standard Arabic are as common as nominal and adjectival loans. Whatever their category, loans are often distinctive in some way (whether because of their syllabic structure, the presence of particular phonemes or their morphological template), since the lexeme usually (though not always) has the same form in both the recipient language and the source language. Examples of loans without any distinctive features are *baṛṛaṛ* 'to justify', and *ðahbi* 'golden'.

A certain number of Standard Arabic verbs with the infix *-t-* or the prefix *sta-* are borrowed, but these verbal patterns can be found elsewhere in Ḥassāniyya.

Certain lexical fields exhibit a particularly high degree of loans from Standard Arabic: anything connected with Islamic studies or abstract concepts (religion, rights, morality, feelings etc.) and, more recently, politics, media and modern material culture. These regularly retain the meaning (or one of the meanings) of the source-language item.

3.4.1.2 Loanwords from Berber

There many lexical items that are probable loans from Berber, with a number of certain cases among them.

Here we may point to several non-Arabic-origin verbs with cognates across a wide range of Berber languages, such as *kṛaṭ* 'to scrape off' (Zenaga *yugṛaḍ*); *šayḍaḍ* 'to make a lactating camel adopt an orphaned calf from another mother' (Zenaga *yaṣṣuḍaḍ* 'to breastfeed', *yuḍḍaḍ* 'to suckle'); *santa* 'to begin' (Zenaga *yassanta* 'to begin', Tuareg *ent* 'to be started, to begin'); *gaymar* 'to hunt from a distance' (Berber *gmər* 'to hunt').

Other verbs are derived from nouns loaned from Berber. Hence, *ɣawba* 'to restrain a camel, put it in an *aɣāba*' (Tuareg *aɣaba* 'jaws'). Sometimes there is both a verb and an adjective stemming from a loaned root, as in *gaylal* 'to have the tail cut' and *agīlāl* 'having a cut tail' (Tuareg *gilel* and *agilal*).

Some loaned Ḥassāniyya nouns are found with the same root (or an equivalent root) in Berber languages other than Zenaga. For example: *agayš* 'male bustard' (Tuareg *gayəs*); *āškəṛ* 'partridge' (Kabyle *tasekkurt* in the feminine form); *tayffārət* 'fetlock (camel)' (Zenaga *tiʔffart*, Tuareg *téffart*); *azāɣər* 'wooden mat ceiling between beams' (Zenaga *azaɣri* 'lintel, beam (of a well)', Tuareg *əzgər* 'to

cross', *ăzəgər* 'crossbeam'); *talawmāyət* 'dew' (Zenaga *tayaṃut*, Tuareg *tălămut*); *(n)tūrža* '*Calotropis procera*' (Zenaga *turžah*, Tuareg *tərza*).

Most of the loanwords cited above are attested in Zenaga (sometimes in a more innovative form than is found in other Berber varieties, such as *yaggīyyay* 'to have a cut tail' where /y/ < *l). However, there are numerous cases where a corresponding Berber item is attested only in Zenaga. In such cases it is difficult to precisely identify the source language, even if the phonology and/or morphology seems to indicate a non-Arabic origin.

Loanwords from Berber seem to be particularly common in the lexicon of fauna, flora, and diseases, as well as in the field of traditional material culture (objects, culinary traditions, farming practices etc.; Taine-Cheikh 2010a; 2014). Unlike the form of the loans, which is often quite divergent from that of the source items, their semantics tends to remain largely unchanged. However, there are some exceptions, notably when the verbs have a general meaning in Berber (cf. above 'to breastfeed' vs. 'to make a lactating camel adopt an orphaned calf from another mother').

3.4.1.3 Loanwords from Sahel languages

Rather few Ḥassāniyya lexical items seem to be borrowed directly from African languages, and the origin of those that are is rarely known precisely. We may note, however, in addition to *gad^{y}* 'dried fish' (< Wolof) and *d^{y}əngra* 'warehouse' (< Soninke), a few terms which appear to be borrowed from Pulaar: *ḅamḅa* 'to carry a child on one's back', *t^{y}əhli* 'roof on pillars' and *kīri* 'boundary between two fields'.

In some regions we find a concentration of loans in particular domains in relation to specific contact languages. For example, in the ancient town of Tichitt, we find borrowings from Azer and Soninke (Jacques-Meunié 1961; Monteil 1939; Diagana 2013): *kā* 'house' (Azer *ka(ny)*, Soninke *ká*) in *kā n laqqe* 'entrance of the house'; *killen* 'path' (Azer *kille*, Soninke *kìllé*); *kunyu ~ kenyen* 'cooking' (Azer *knu ~ kenyu*, Soninke *kìnŋú*).

A significant list of loanwords from Songhay has been compiled by Heath (2004) in Mali, including e.g.: *ṣawṣab* (< *sosom ~ sosob*) 'pound (millet) in mortar to remove bran from grains'; *daydi ~ dayday* (< *deydey*) 'daily grocery purchase'; *ākāṛāy* (< *kaarey*) 'crocodile'; *sari* (< *seri*) 'millet porridge'. Only *sari* has been recorded elsewhere in Mauritania (in the eastern town of Walata). On the other hand, all authors who have done field work on the Ḥassāniyya of Mali (particularly in the region of Timbuktu and the Azawad), have noted loanwords from Songhay. This is true also of Clauzel (1960) who, as well as a number of Berber

loanwords, gives a small list of Songhay-derived items used in the salt mine of Tāwdenni, such as *titi* 'cylinder of saliferous clay used as a seat by the miners' (< *tita*) and *tʸar* 'adze' (< *tʸara*).

3.4.1.4 Loanwords from Indo-European languages

The use of loanwords from European languages tends to vary over time. Thus, a large proportion of the French loanwords borrowed during the colonial period have more recently gone out of use, such as *bəṛṭmāla* or *qoṛṭmāl* 'wallet' (< *porte-monnaie*), *dabbīš* 'telegram' (< *dépêche* 'dispatch') or *ṣaṛwaṣ* 'to be very close to the colonizers' (< *service* 'service'). This is true not only of items referring to obsolete concepts (such as the currency terms *sūvāya* 'sou' or *ftən/vəvtən* 'cent', likely < *fifteen*), but also of those referring to still-current concepts which are, however, now referred to with a term drawn from Standard Arabic (e.g. *minəstr* 'minister', replaced by *wazīr*). This does not, however, eliminate the permanence of some old loanwords such as *wata* 'car' (< *voiture*) or *maṛṣa* 'market' (< *marché*).[6]

Although not unique to Ḥassāniyya, the frequency of the emphatic phonemes (especially /ṣ/ and /ṭ/) in loans from European languages is notable. Consider, in addition to the treatment of *service*, *porte-monnaie* and *marché* as noted above, that of *baṭṛūn* 'boss' (< *patron*), which gives rise to *tbaṭṛan* 'to be(come) a boss' *ṭawn* 'ton' (< *tonne*).

3.4.2 More complex cases

3.4.2.1 *Wanderwörter*

Various Arabic lexical items derive from Latin, Armenian, Turkish, Persian, and so on. In the case of, for example, the names of calendar months, or of items such as trousers (*sərwāl*), these terms are not borrowed directly from the source language by Ḥassāniyya and are found elsewhere (e.g. *balbūza* 'eyeball' < Latin *bulbus*, attested throughout the Maghreb). The history of such items will not be dealt with here. We can, however, mention the case of some well-attested terms in Ḥassāniyya that appear to have been borrowed from sub-Saharan Africa.

One such is *māṛu* 'rice', which seems to come from Soninke (*máarò*), although it is also attested in Wolof (*maalo*) and Zenaga (*mārih*). Another term, which is just as emblematic, is *mbūṛu* 'bread', whose origin has variously been attributed to Wolof, Azer, Mandigo, and even English *bread*.

[6] Ould Mohamed Baba (2003) gives an extensive list of loanwords from French and offers a classification by semantic field.

To these very everyday terms, we may also add *ṃutri* 'pearl millet' and *makka* 'maize', which have the same form both in Ḥassāniyya and in Zenaga. The first is a loanword from Pulaar (*muutiri*). The second is attested in many languages and seems to have come from the placename Mecca.

As for *garta* 'peanut', *ḷāḷo* ~ *ḷaḷu* 'pounded baobab leaves that serve as a condiment' (synonym of *taqya* in the southwest of Mauritania) and *kəddu* 'spoon', these appear to be used just as frequently in Pulaar as they are in Wolof.

3.4.2.2 Berberized items

Despite the absence of any Berber affixes in the loanwords listed in §3.4.2.1, only *kəddu* 'spoon' is regularly used with the definite article. In this regard, these loanwords act like words borrowed from Berber, or more generally, those with Berber affixes.

It is, in fact, difficult to prove that a noun with this kind of affix is definitely of Berber origin, since we find nouns of various origins with Berber affixes. Some of them are loanwords from the languages of the sedentary people of the valley, such as *adabāy* 'village of former sedentary slaves (*ḥṛāṭīn*)' (< Soninke *dèbé* 'village'); *iggīw* ~ *īggīw* 'griot' (Zenaga *iggiwi*, borrowed from Wolof *geewel* or from Pulaar *gawlo*). Others are borrowed from French: *agāṛāž* 'garage'; *təmbīskit* 'biscuit'. Even terms of Arabic origin are Berberized, as is likely the case with with *tasūvra* 'large decorated leather bag for travelling' (cf. *sāvər* 'to travel') or *tāẓəẓmīt* 'asthma' (cf. CA *zağma* 'shortness of breath when giving birth').

3.4.2.3 Reborrowings

Instances of back and forth between two languages – primarily Ḥassāniyya and Zenaga – seem to be the reason for another type of mixed form, illustrated previously in §3.2.2.1 by the Zenaga verbs *yassəðbah* 'to leave in the afternoon' and *yišnar* 'to orient oneself'.

Ḥassāniyya *sayɣnan* 'to mix gum with water to make ink' provides another example, where this time the points of departure and arrival seem to be from the Arabic side. In fact, this loanword is a borrowing of Zenaga *yassuɣnan* 'to thicken (ink) by adding gum', a verb formed from *əssaɣan* 'gum'. This noun in turn appears to be an adaptation of the Arabic *samɣa* 'ink'.

In the case of *sla* 'placenta', there is a double round-trip between the two languages, this time without metathesis: after a passage from Arabic to Zenaga (> *əs(s)la*), there is return to Ḥassāniyya with the causative verb *sasla* 'to soak a

hide', and a second loan into Zenaga with the reflexive form *(yə)stasla* 'to start to lose fur (of soaked hides)'.

3.4.2.4 Calques

Calques are undoubtedly common, but they are particularly frequent in locutions such as *rəggət əž-žəll* 'susceptibility' and *bū-damʕa* 'rinderpest' (literally 'thinness of skin' and 'the one with a tear'). These are exact calques of their Zenaga equivalents *taššəddi-n əyim* and *ən-anḍi* (Taine-Cheikh 2008a).

3.4.2.5 Individual variation

Receptivity to loanwords differs from one individual to another. This is natural when we are dealing with bilingual speakers and this probably explains the special features of the Ḥassāniyya of the Awlād Banʸūg (often bilingual speakers of Ḥassāniyya and Wolof) or the Ḥassāniyya of Mali (where Arabic speakers often speak Songhay and sometimes Tamasheq). However, it also depends on the individuals in question in terms of what we might call their "loyalty" to the language, whether the language is under pressure from Moroccan Arabic koiné in Morocco (Taine-Cheikh 1997b; Heath 2002; Paciotti 2017), or whether it is imposed as a lingua franca in Mauritania (Dia 2007).

4 Conclusion

The principal domain affected by contact in Ḥassāniyya is that of the lexicon (though an assessment in percentage terms is not at present possible). However, the integration of loanwords – in particular those from Standard Arabic and Berber – has resulted in a significant enrichment of the phonological system and of the inventory of nominal patterns. The effects of contact on the verbal morphology and syntax of the dialect are more indirect. The major developments in Ḥassāniyya seem most likely to instead be a product of internal evolution. In certain cases, Zenaga has probably had an influence; in others, we rather witness instances of parallel evolution.

In future, by studying the vehicular Ḥassāniyya of Mauritania and of the border regions (southern Morocco, southern Algeria, Senegal, Niger, and so on) we will perhaps discover new developments as a result of contacts triggered by the political and societal changes of the twenty-first century.

Further reading

Links between Ḥassāniyya and other languages are particularly complex at the level of semantics and lexicon. On these topics, beyond the available Ḥassāniyya and Zenaga dictionaries (Heath 2004; Taine-Cheikh 1988–1998; 2008b), readers may consult the available studies of specific fields (Monteil 1952; Taine-Cheikh 2013) or particular templates (Taine-Cheikh 2018a).

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	M	masculine
CA	Classical Arabic	PL	plural
DEF	definite article	REL	relativizer
Hass.	Ḥassāniyya Arabic	SG	singular
IMPF	imperfect	Zen.	Zenaga

References

Aguadé, Jordi. 1998. Relatos en hassaniyya recogidos en Mħāmīd (Valle del Dra, sur de Marruecos). *Estudios de Dialectología Norteafricana y Andalusí* 3. 203–215.

Cantineau, Jean. 1960. *Etudes de linguistique arabe*. Paris: Klincksieck.

Clauzel, Jean. 1960. *L'exploitation des salines de Taoudenni*. Institut de recherches sahariennes de l'Université d'Alger: Imprimerie Protat Frères.

Cohen, David. 1963. *Le dialecte arabe ḥassānīya de Mauritanie*. Paris: Klincksieck.

Cohen, Marcel. 1924. *Le système verbal sémitique et l'expression du temps*. Paris: Ernest Leroux.

Dia, Alassane. 2007. Uses and attitudes towards Hassaniyya among Nouakchott's Negro-Mauritanian population. In Catherine Miller, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.), *Arabic in the city: Issues in dialect contact and language variation*, 325–344. London: Routledge.

Diagana, Ousmane. 2013. *Dictionnaire soninké-français (Mauritanie)*. Paris: Karthala.

Heath, Jeffrey. 2002. *Jewish and Muslim dialects of Moroccan Arabic*. London: Routledge.

Heath, Jeffrey. 2004. *Hassaniya Arabic (Mali)–English–French dictionary*. Wiesbaden: Harrassowitz.

Jacques-Meunié, Djinn. 1961. *Cités anciennes de Mauritanie*. Paris: Klincksieck.

Monteil, Charles. 1939. La langue azer. In Théodore Monod (ed.), *Contributions à l'étude du Sahara occidental*, 213–341. Paris: Larose.

Monteil, Vincent-Mansour. 1952. *Essai sur le chameau au Sahara Occidental*. Saint-Louis, Senegal: Centre IFAN–Mauritanie.

Ould Mohamed Baba, Ahmed-Salem. 2003. Emprunts du dialecte ḥassāniyya à la langue française. In Ignacio Ferrando & Juan José Sanchez Sandoval (eds.), *AIDA 5th conference proceedings, Cádiz, September 2002*, 61–74. Cadiz: Universidad de Cádiz, Servicio de Publicaciones.

Paciotti, Luca. 2017. *Conflitti linguistici e glottopolitica in Marocco: La situazione della ḥassāniyya nella regione di Guelmime-Oued Noun*. Naples: Università degli Studi di Napoli "L'Orientale". (Master's dissertation).

Taine-Cheikh, Catherine. 1988–1998. *Dictionnaire ḥassāniyya–français*. Paris: Geuthner.

Taine-Cheikh, Catherine. 1997a. Les emprunts au berbère zénaga: Un sous-système vocalique du ḥassāniyya. *Matériaux Arabes et Sudarabiques (GELLAS)* 8. 93–142.

Taine-Cheikh, Catherine. 1997b. Les hassanophones du Maroc: Entre affirmation de soi et auto-reniement. *Peuples méditerranéens* 79. 85–102.

Taine-Cheikh, Catherine. 2003. Les valeurs du préfixe *s-* en hassaniyya et les conditions de sa grammaticalisation. In Ignacio Ferrando & Juan José Sanchez Sandoval (eds.), *AIDA 5th conference proceedings, Cádiz, September 2002*, 103–118. Cadiz: Universidad de Cádiz, Servicio de Publicaciones.

Taine-Cheikh, Catherine. 2004. Le(s) futur(s) en arabe: Réflexions pour une typologie. *Estudios de Dialectología Norteafricana y Andalusí* 8. 215–238.

Taine-Cheikh, Catherine. 2008a. 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. 2008b. *Dictionnaire zénaga–français: Le berbère de Mauritanie par racines dans une perspective comparative*. Cologne: Rüdiger Köppe.

Taine-Cheikh, Catherine. 2009. Les morphèmes de futur en arabe et en berbère: Réflexions pour une typologie. *Faits de Langue* 33. 91–102.

Taine-Cheikh, Catherine. 2010a. Aux origines de la culture matérielle des nomades de Mauritanie: Réflexions à partir des lexiques arabes et berbères. *The Maghreb Review* 35(1–2). 64–88.

Taine-Cheikh, Catherine. 2010b. The role of the Berber deictic and TAM markers in dependent clauses in Zenaga. In Isabelle Bril (ed.), *Clause linking and clause hierarchy: Syntax and pragmatics*, 355–398. Amsterdam: John Benjamins.

Taine-Cheikh, Catherine. 2013. Des ethnies chimériques aux langues fantômes: L'exemple des Imraguen et des Nemâdi de Mauritanie. In Carole de Féral (ed.), *In and out of Africa: Languages in question. In honour of Robert Nicolaï*, vol. 1: *Language contact and epistemological issues*, 137–164. Leuven: Peeters.

Taine-Cheikh, Catherine. 2014. Les voies lactées: Le lait dans l'alimentation des nomades de Mauritanie. *Awal – Cahiers d'études berbères* 42. 27–50.

Taine-Cheikh, Catherine. 2016. *Etudes de linguistique ouest-saharienne*. Vol. 1: *Sociolinguistique de l'aire hassanophone*. Rabat: Centre des Etudes Sahariennes.

Taine-Cheikh, Catherine. 2018a. Ḥassāniyya Arabic in contact with Berber: The case of quadriliteral verbs. In Stefano Manfredi & Mauro Tosco (eds.), *Arabic in contact*, 135–159. Amsterdam: John Benjamins.

Taine-Cheikh, Catherine. 2018b. Historical and typological approaches to Mauritanian and West Saharan Arabic. In Clive Holes (ed.), *Arabic historical dialectology: Linguistic and sociolinguistic approaches*, 293–315. Oxford: Oxford University Press.

Part II

Language change through contact with Arabic

Chapter 6

Neo-Aramaic

Eleanor Coghill
Uppsala University

This paper examines the impact of Arabic on the North-Eastern Neo-Aramaic dialects, a diverse group of Semitic language varieties native to a region spanning Iraq, Turkey, Syria and Iran. While the greatest contact influence comes from varieties of Kurdish, Arabic has also had considerable influence, both directly and indirectly via other regional languages. Influence is most apparent in lexicon and phonology, but also surfaces in morphology and syntax.

1 Current state and historical development

The Aramaic language (Semitic, Afroasiatic) has nearly three thousand years of documented history up to the present day. Once widely used, both as a first language and as a language of trade and officialdom, since the Arab conquests of the seventh century it has steadily shrunk in its geographical coverage. Today its descendants, the Neo-Aramaic dialects, only remain in pockets, especially in remoter regions, and are spoken almost exclusively by religious–ethnic minorities. Four branches of the language family exist today: due to diversification these cannot be considered a single language. Indeed, the largest branch, North-Eastern Neo-Aramaic (NENA), which is treated in this chapter, itself consists of many mutually incomprehensible dialects. Its closest relation is Ṭuroyo/Ṣurayt, which is spoken by Christians, known as *Suryoye*, indigenous to the area immediately west of NENA's western edge in Turkey. Another member of this branch (Central Neo-Aramaic) was Mlaḥso, but this was nearly wiped out during the First World War, and its last speaker apparently died in the 1990s.

The NENA dialects are, or were, spoken in a contiguous region stretching across northeastern Iraq, southeastern Turkey, northeastern Syria and northwestern Iran. The majority ethnicity in this region is the Kurds. NENA's native

Eleanor Coghill. 2020. Neo-Aramaic. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 127–156. Berlin: Language Science Press. DOI:??

speakers are exclusively from Christian and Jewish communities. The Christians belong to a variety of churches: the Church of the East, the Chaldean Catholic Church (which split off from the Church of the East when it came into communion with Rome), and (in fewer numbers) the Syriac Orthodox Church and its uniate counterpart, the Syriac Catholic Church. The Christians' traditional religious–ethnic endonym is *Surāye* and they call their language *Sūraθ* or *Sūrət* (depending on dialectal pronunciation). In other languages, and sometimes in their own, they identify mainly as Assyrians or Chaldeans.

The Jews are called *hudāye* or *hulāʔe* (depending on dialectal pronunciation), and they call their language *lišāna deni/nošan* 'our language' or *hulaula* 'Jewishness'. In Israel, where most now live, they are known as *kurdím*, reflecting their geographical origin in the Kurdish region, rather than their ethnic identity.

Historically, the NENA-speaking Christians usually lived in rural mono-ethnic villages and predominantly practiced agriculture, animal husbandry and crafts. Jews lived in both villages and towns, alongside other ethnic groups such as Kurds. They had diverse professions: tradesmen (pedlars, merchants and shopkeepers), craftsmen, peasants and landowners (Brauer & Patai 1993: 205, 212).

The region to which NENA is indigenous was, until, the twentieth century, highly diverse in terms of ethnicity, religion and language. Some of this diversity remains, but a great deal has been lost, due to the persecutions and ethnic cleansing that went on during that century and which were not unknown prior to it. During the First World War, Christian communities in Anatolia, being viewed as a fifth column in league with Russia, suffered murderous attacks and deportations. This affected not only Armenians and Greeks, but also the Sūraθ-speaking Surāye and Ṭuroyo-speaking Suryoye, as well as the many Arabic-speaking Christian communities in the region (the extirpation of some of these is documented in Jastrow 1978: 3–17).[1] By the 1920s, the Hakkari province of Turkey had been emptied of its many communities of Surāye: survivors ended up in Iraq and Iran. Some Sūraθ-speaking villages remained in the neighbouring Şırnak and Siirt provinces, but in the late twentieth century these too were mostly emptied of their inhabitants, during the conflict between the Turkish state and the Kurds.

In Iraq too the twentieth century was far from peaceful for the NENA-speaking communities. After a massacre in the 1930s, a proportion of the survivors of the genocide moved from Iraq to Syria, where they settled along the Khabur river, still in their tribal groups. Others remained in Iraq, in some places in their original

[1]The relationship between language and ethno-religious identity was and remains complex. Many Christians belonging to the Syriac churches spoke and continue to speak yet other regional languages, including varieties of Turkish, Armenian and Kurdish.

communities, in other places in mixed communities, where a koiné form of Sūraθ arose. After the founding of Israel, there was a backlash against Jews in Iraq, and almost all Jews left the country for Israel during the 1950s. In Israel their heritage and language were for the most part not appreciated and the language was not passed on to younger generations. Most remaining speakers are now elderly and some dialects have already died out.

From the 1960s onwards, conflicts between Kurdish groups and the Iraqi state resulted in the destruction of numerous northern Iraqi villages, including many Christian ones. Other villages were appropriated by Kurdish tribes. The war in 1990–1991, the international sanctions and the invasion of 2003 and subsequent instability further affected these communities, as they did all Iraqis, and resulted in a dramatic shrinking of the Christian community in Iraq. In 2014, when ISIS captured large swathes of northern Iraq, many Christians and other non-Sunni minorities had to leave their villages overnight. These villages were later recaptured, but, in the absence of extensive rebuilding and due to fears of a recurrence, many inhabitants have not returned and seek to leave the country. The outlook is therefore bleak for these communities and for their language.

2 Contact languages

The main contact language for NENA is – and has been for long time – Kurdish (Iranian, Indo-European), in its many varieties, as Kurds are by far the largest ethnic group in the region as a whole, excepting Iranian Azerbaijan, where Azeris predominate.[2] Kurds have also been politically dominant: during the Ottoman period, Christians and Jews were in the power and under the protection of local Kurdish rulers, the aghas (see Sinha 2000: 11–12; Brauer & Patai 1993: 223). Most NENA speakers in the Kurdish-speaking areas at this time seem to have spoken the local Kurdish dialect.[3] It is not surprising, therefore, that there is more influence from Kurdish than from any other language across most if not all of the NENA dialects, even if its extent varies from dialect to dialect.

What role, then, has Arabic played? To summarize: there has been longstanding direct contact with small Arabic-speaking communities in what are other-

[2]Small communities of Turkic-speaking Turkmans are also found within northern Iraq. Their dialects share features with both Anatolian Turkish varieties and Iranian Azeri (Bulut 2007).

[3]For such information we rely mainly on statements in grammatical descriptions, where the researcher asked their informants about this. For instance, Hoberman (1989: 9) states, "All my informants who grew to adulthood in Kurdistan report that they spoke fluent Kurdish (Kurmanji)". Other references for Jews' competence in Kurdish are: Sabar (1978: 216), Mutzafi (2004: 5), Khan (2007: 198) and Khan (2009: 11); for the Christians see Sinha (2000: 12–13) and Khan (2008: 18).

wise Kurdish-speaking regions; there has been indirect contact through loans transmitted via Kurdish and Azeri varieties; finally, there has been intense contact more recently due to the establishment of states with Arabic as the national language, as well as various other modern developments. In the remainder of this section, we will go through these three types of contact in turn.

Although the region is not majority Arabic-speaking, there have been long-standing Arabic-speaking communities in certain parts of it: moreover many of these were Jewish and Christian, like the NENA-speakers, so one might well expect more social contacts with them. The Arabic dialects across the region are overwhelmingly of the *qəltu* Mesopotamian–Anatolian type (contrasted with the southern Iraqi/Bedouin *gələt* type).[4]

Christian *qəltu* Arabic speakers could be found in the city of Mosul (alongside *qəltu* Arabic speakers of other religions) on the edge of the NENA-speaking Nineveh Plain (also known as the Mosul Plain). They are also present in two villages on the Nineveh Plain, namely Bəḥzāni and Baḥšiqa. Arabic-speaking Yazidis[5] also live in these villages, as well as (in Baḥšiqa) some Muslim Arabs (Jastrow 1978: 24). The Christian NENA speakers of the Nineveh Plain, therefore, had ample opportunity to come into contact with Arabic. To find more Christian Arabic-speaking communities in or near the NENA region, we have to travel quite far, to what are now the Turkish provinces of Şırnak, Siirt and Mardin. In this region there were many Christian *qəltu* Arabic-speaking communities living in villages and towns until the First World War; fewer afterwards. The settlements with such communities included Āzəḫ (Turkish *İdil*) and Ǧazīra (*Cizre*) in Şırnak province, as well as provincial centres Siirt and Mardin (Jastrow 1978: 1–23). Thus, Christian Arabic speakers were in close proximity to speakers of NENA dialects in the Bohtan and Cudi regions of Şırnak province, as well as to speakers of Ṭuroyo/Ṣurayt in Mardin Province.

Jewish *qəltu* Arabic-speaking communities were also found in both northern Iraq and southeastern Turkey. In Iraq, Arabic was spoken by the Jews of Mosul, ʕAqra (Kurdish *Akre*) and Arbil (Erbil; Kurdish *Hawler*), as well as of the village of Ṣəndor, near Duhok (Hoberman 1989: 9). These all left in the 1950s. Further

[4]The two types of Mesopotamian–Anatolian Arabic dialects are labelled by scholars according to the shibboleth of the form 'I said': *qəltu* vs. *gələt* (Blanc 1964: 5–8). *qəltu* dialects realize *q as /q/, while *gələt* dialects (such as Muslim Baghdadi), which are Bedouin or Bedouin-influenced, realize it as /g/. *Qəltu* dialects also preserve the 1SG inflection *-u* on the suffix-conjugation verb. See Talay (2011) for an overview of Mesopotamian–Anatolian Arabic varieties. Note that some Bedouin influence may be seen in the Muslim *qəltu* dialects spoken on the plain south of Mardin (Jastrow 1978: 30).

[5]Elsewhere, Yazidis are Northern Kurdish-speaking.

afield, there were also some Jewish Arabic speakers in Urfa, Diyarbakır, Siverek and Çermik (Jastrow 1978: 4), who also migrated to Israel. There are known to have been contacts between NENA-speaking and Arabic-speaking Jews, through family connections and commerce. Mutzafi (2004: 6) reports such contacts involving the Jewish men of Koy Sanjaq and the Arabic-speaking Jews of Kurdistan. Sabar (1978: 216–217) relates that the Jews of Zakho would visit relatives who had moved to Mosul and Baghdad. On the other hand, Hoberman (1989: 9) stated that the Jews of ʕAmədya knew no more than a few words of Iraqi Arabic.

To sum up, historically, Christian NENA speakers only had direct local contact with Arabic speakers (of their own faith) in Mosul and the Nineveh Plain in Iraq and Şırnak province in Turkey. The NENA-speaking Jews, on the other hand, had Arabic-speaking co-religionists not only in Mosul, but also within Iraqi Kurdistan itself.

While most NENA dialects show greatest influence from the majority languages of the region – Kurdish and (in Iranian Azerbaijan) Iranian Azeri – these also played a role in transferring Arabic influence to NENA. Arabic, as the language of Islam, has had a great influence on Kurdish varieties and Azeri, especially in the lexicon, and many originally Arabic words have been transmitted to NENA via these languages. Sometimes it is difficult to identify the immediate donor of such words, but phonetics and morphology can help (see §3.1.1).

During the twentieth century, with the founding of the states of Iraq and Syria, Arabic became the language of the states that most NENA-speakers found themselves in. They came into contact with it through education, officialdom, military service, radio and trade. Many Christians from the north of Iraq moved south to the major (Arabic-speaking) cities, Mosul, Baghdad and Baṣra, where, in some cases, they shifted to speaking Arabic, while keeping in close contact with relatives back in the north. By the end of the twentieth century most NENA speakers in Iraq and Syria would have been at ease in Arabic. Naturally these later developments did not affect speakers in Turkey and Iran, who, instead, developed greater competence in Turkish and Persian, respectively. Jewish speakers from Iraq, who had left the region by the end of the 1950s, would have had less exposure to Arabic through these means.

It should be mentioned that there has also been influence from European languages, namely from French (via the influence of the Catholic Church among the Chaldean Catholic communities) and from English (dating to the British Mandate period, as well as the period of globalization from the late twentieth century), though some lexical borrowings from these languages may have been mediated by Arabic.

3 Contact-induced changes in North-Eastern Neo-Aramaic

Contact influence on NENA[6] seems to have arisen mainly through long-term bi- and multi-lingualism, rather than language shift. Indeed, if any shift has taken place, it is more likely to have involved NENA speakers who converted to Islam and shifted to Kurdish.[7] Furthermore, much of Iraq was in earlier times Aramaic-speaking, so it can be assumed that over the centuries a shift took place from Aramaic to Arabic. Some Aramaic substrate features can indeed be seen in Iraqi Arabic dialects, such as a kind of differential object marking (Coghill 2014: 360–361).

Using Van Coetsem's (1988; 2000) distinctions between changes due to borrowing (by agents dominant in the recipient language) and imposition (by agents dominant in the source language), the contact influences from Arabic attested in NENA are clearly of the first kind, namely borrowing.

Borrowing from Arabic into NENA is of interest particularly as a case of transfer between related and typologically similar languages, as both are Semitic. Like Arabic and other Semitic languages, NENA has in its verbal morphology, and to a lesser extent in its nominal morphology, a non-concatenative root-and-pattern system, complemented by affixes. Thus, with the triradical root *√šql*, we get such

[6]Sources for the main contact languages, if not indicated, are as follows: Iraqi Arabic (specifically Muslim Baghdadi): Woodhead & Beene (1967); Northern Kurdish (i.e. Kurmanji/Bahdini): Chyet (2003). Although Muslim Baghdadi Arabic is not the dialect in closest contact with NENA, as a Mesopotamian dialect it shares much lexicon with more northerly varieties (which do not have a dictionary). The transcription of Northern Kurdish words is based on the conventional orthography, as given in Chyet (2003: xxxix–xl): an IPA transcription is also given. The source for the Christian Alqosh and Christian Telkepe data is the author's own fieldwork. Other sources are referenced in the text. The author's own NENA data is transcribed in IPA except as follows: *č* [ʧ], *j* [ʤ] (equivalent to Arabic *ǧ*), *y* [j], *ḥ* [ħ], *x* between [x] and [χ], and *ġ* between [ɣ] and [ʁ]. Apart from *ḥ*, consonants with a dot under are the emphatic (velarized/pharyngealized) versions of the undotted consonant; for instance, the symbol *ð̣* represents [ðˤ]. Some dialects have emphasis extending across whole words: such words are conventionally indicated with a superscript cross, e.g. ˣ*sadra* (equivalent to *ṣaḍṛa*). The schwa symbol *ə* is used to transcribe a NENA vowel that is, in non-emphatic contexts, typically pronounced as [ɪ]. Phonemically contrastive length in vowels is indicated with a macron, e.g. *ā* [aː]. The vowels /i/, /e/ and /o/ are usually realized long: [iː], [eː] and [oː]. NENA words from other sources have had their transcription adjusted in some cases to bring them closer to this system: the original transcription may be checked in the referenced sources.

[7]It often happened that Christian girls were (occasionally by arrangement, but often unwillingly) kidnapped by Kurds for the purpose of marriage. Any children would have been considered Kurds.

forms as *k-šāqəl* 'he takes', *šqəl-lə* 'he took', *šqāla* 'taking', *šaqāla* 'taker', *šqila* 'taken', and so on.

Arabic influence in NENA is considerable in the realm of the lexicon, but this has very often occurred via other contact languages, rather than directly. (All the contact languages show great influence from Arabic, at least in the lexicon). Direct lexical borrowing or morphological and structural borrowing from Arabic are less common: they are however well attested in the Christian dialects of the Nineveh Plain, as well as some Jewish dialects of the Lišāna Deni branch in northern Iraq, including the dialects of Zakho, Nerwa and ʕAmədya (Kurdish *Amêdî*, Arabic *al-ʕAmādiyya*).

It is difficult to establish with any certainty which contact influences entered the dialects at which time. The earliest Christian and Jewish NENA texts (from the 16th and 17th centuries)[8] already show considerable contact influence from Kurdish and Arabic. The extent of Arabic influence in the early Jewish Lišāna Deni texts (Sabar 1984) is quite surprising. The towns in which these texts originate lie deep in Kurdistan, relatively far from the Arabic speaking part of Iraq. As we have seen in §2, however, Jews in Kurdistan had contacts with Arabic-speaking co-religionists. Some contact influence in the NENA dialects is clearly of recent date, such as loanwords from English, which probably date to the twentieth century. The prospective construction of the Christian Nineveh Plain dialects, which appears to be a structural borrowing from vernacular Arabic (see §3.4), seems to have developed only in the last hundred years or so (Coghill 2010: 375).

By the end of the twentieth century, Arabic was having an immense influence on the speech of Christian Aramaic-speaking communities living in northern Iraq, expecially those close to Mosul, such as the town of Qaraqosh. Khan (2002: 9) found that most people from Qaraqosh introduced Arabic words and phrases into their Neo-Aramaic without adaptation. Khan attributes this to the policy of Arabicization in Iraq, which meant that schoolchildren were only educated in Arabic. He found significantly greater influence from Arabic in the younger generation's speech. In Christian Qaraqosh, as in the neighbouring dialects of Christian Alqosh and Christian Telkepe (author's fieldwork), a large number of Arabic loanwords have recently been absorbed into the lexicon. Nevertheless, as Khan remarks, "the proportion of Arabic loans that have penetrated the core vocabulary of the dialect and replaced existing Aramaic words are relatively few." This may, however, not be the case with speakers who have grown up in Arab-

[8]The Jewish manuscripts date to the 17th century, but the texts may have been composed earlier (Sabar 1976: xxix, xliii–xlvi). The Christian manuscripts date to the 18th century but the composition of the texts can be dated to the 16th and 17th centuries (Mengozzi 2002: 16).

majority cities such as Baghdad. In my admittedly limited experience with such speakers, they use a noticeably greater proportion of Arabic loanwords, even sometimes for basic vocabulary, e.g. Iraqi Arabic *ðēʕa* for *māθɒ* 'village' (heard from a Christian Telkepe speaker who grew up in Baghdad before settling in the US).

3.1 Lexicon

3.1.1 Introduction

All NENA dialects have adopted a large number of loanwords. While Kurdish predominates among these, Arabic loanwords are also common, especially among the Christian dialects of the Nineveh Plain and the Jewish Lišāna Deni dialects.

Khan (2002: 516) makes a useful distinction for Christian Qaraqosh between "(i) loan-words that do not have any existing Aramaic equivalent and (ii) those for which a native Aramaic substitute is still available in the dialect."[9] These two types seem to reflect two layers of borrowing, an earlier one and a recent one, which, in many cases, is akin to code-switching. Most Kurdish loans belong to the first type, while Arabic loans are most common in the second, though earlier loans do exist. Borrowed Arabic nouns of the second type show little or no adaptation to native morphology, Khan finds. Verbs, however, are always adapted to NENA verbal morphology. Most are slotted into the existing NENA verbal derivations (see §3.1.4).

Khan (2002: 516) remarks that speakers of Christian Qaraqosh are generally aware of the Aramaic alternatives to these Arabic loans and can give them if asked. It could be, however, that subsequent generations will have had little exposure to the older synonyms.[10] Khan notes that some of these older synonyms are themselves loanwords, in some cases from Arabic, but so integrated and long-standing that many speakers may not be aware of this. Examples include the recent Arabic loan *fəkr* (< Arabic *fikr*) and the older loan *taxmanta* (f. infinitive of NENA *√txmn* Q 'to think', denominal < Arabic *taḫmīn* 'estimation'; see §3.1.4), both meaning 'thought'.

Many loanwords are common to several languages of the region, especially words specific to local culture or to technologies. While the ultimate source can usually be identified, it can sometimes be hard to determine the immediate donor

[9]Note, however, that apparent synonyms are not always identical in meaning. Christian Alqosh *šəbbakiyə* (< Ar. *šubbāk*) is used for a modern glass window, while the inherited lexeme *kāwə* is used for the traditional type of window.

[10]The fieldwork for the monograph on this dialect was carried out around the year 2000.

of the loan.

Nevertheless, there is sometimes evidence that can establish the immediate donor. This is the case, for example, for Arabic words ending in the feminine suffix *tāʔ marbūṭa* (Standard Arabic *-a(t)*). The Arabic morpheme is realized with the final /t/ in suffixed forms and in the construct (i.e. followed by a possessor). When borrowed into NENA, the /t/ is not realized in the absolute (isolated) form of the word, as in Arabic, e.g. Alqosh *sāʕa* 'hour' (Ar. *sāʕa*). This contrasts with Kurdish, which has the /t/ in all forms, e.g. N. Kurd. *saʿet* [sɑːˈʕæt] 'hour'. In some NENA dialects, in certain words, the /t/ appears as *-ət-* in suffixed forms, replicating a pattern in (*qəltu*) Arabic. Sometimes this leads to back-formations (see §3.3.1). In other items the *tāʔ marbūṭa* is realized as *-at* in all contexts, as it typically is in Kurdish, and this suggests it was borrowed via Kurdish. An example of the latter is Jewish Betanure/Jewish Challa *ʕaširat* 'tribe', pl. *ʕaširatte* (Mutzafi 2008: 103; Fassberg 2010: 270). This is borrowed from Northern Kurdish *ʿeşîret* [ʕæʃiːˈræt], which borrowed it from Ar. *ʕašīra(t)* 'tribe', almost certainly via Persian and/or Ottoman Turkish. Another example, *ʕādat* 'custom', is given by Maclean in his grammar of "Vernacular Syriac" (Maclean 1895: 35), where he states that nouns ending in *-at* are feminine.[11] Fox (2009: 91), writing of Christian Bohtan, also views Arabic loans ending in *-at* as having been borrowed via Kurdish. Examples in this dialect are: *sahat* 'hour', *hakowat* 'tale', *qəṣṣat* 'story', *kəflat* 'family' (< N. Kurd. *kuflet* [kʊfˈlæt] ~ *kʾulfet* [kʰʊlˈfæt] 'wife, family' < Ar. *kulfa* 'trouble') and *məllat* 'nation' (< N. Kurd. *milet* [mɪˈlæt] < Ar. *milla*). Some of the same examples (*məḷḷat* and *qəṣṣat*) may also be found in Christian ʕUmra: Hobrack (2000: 108) takes these to have been borrowed via Turkish, but, given the overwhelming influence of Kurdish in the region, it seems more plausible that they were borrowed via Kurdish.[12]

Sometimes there are other indications in the word's form that it was borrowed via Kurdish: the common NENA word *šūla* 'work' derives ultimately from Arabic

[11]In Maclean's dictionary (Maclean 1901: 235), he gives *ʕādat* (orthography adjusted) as the form in the Christian Urmia dialect and as one of the variants in "Alqosh", by which he means the Nineveh Plain dialects (the other variant being *ʕāde*, which, lacking the final /t/, appears to be directly borrowed from Arabic). He gives *ʕādəta,* on the other hand, for his "Ashirat" dialect group, which was spoken in "central Kurdistan" (today's Hakkari province of Turkey). This looks like the back-formations from direct Arabic loans discussed in §3.3.1, which is a little surprising, as one would not expect much direct contact with Arabic in that region. It is, however, a large and diverse group of dialects, and he does not specify in which precise dialect it was attested.

[12]The Kurdish forms attested in dictionaries are not always what we would expect as the sources of these forms, however. Thus we find *ħekyat* [ħækjɑːt] ~ *ħikyet* [ħɪkˈjæt] 'story' and *qise* [qɪˈsæ] 'story' (not *qiset*). A variant of the latter ending in /t/, however, is found in a nineteenth-century dictionary cited in Chyet (2003: 490–491).

šuɣl. Northern Kurdish has also borrowed this word, as *şuxul* [ʃuˈxul] with a variant *şûl* [ʃuːl]. It is perhaps the latter which is the immediate origin of the NENA word.

The gender in NENA can also suggest the immediate source of a loanword. For instance, *qalam* 'pen' in Arabic has masculine gender, but, loaned into Northern Kurdish as *qelem*, it may have feminine or masculine gender (Chyet 2003: 478; Rizgar 1993: 322). That *qalāma* 'pen' has feminine gender in certain NENA dialects (e.g. Alqosh; Coghill 2004: 199) suggests that it was borrowed via Kurdish, not directly from Arabic.

It is difficult to date loanwords in a predominantly unwritten language. Nevertheless, we do have written texts in both the Christian Nineveh Plain and the Jewish Lišāna Deni dialects going back at least four hundred years, and even in early texts the proportion of lexemes that were borrowed was high. Arabic loans are conspicuous in both sets of texts. Sabar (1984: 208) found that in a typical Jewish text from Nerwa, 30% of lexemes are ultimately of Arabic origin (whether directly or via another language).

Loanwords may be adapted to varying degrees and in varying ways to the recipient language. §§3.1.2–3.1.5 deal with the ways in which loans in different word classes may be integrated, as well as the ways in which they retain characteristics of the donor language, focusing on Arabic loans.

3.1.2 Integration of nouns

Most NENA nouns end in the nominal suffix *-a* (usually, but not exclusively, masculine nouns) or *-ta~-θa* (feminine nouns). Older borrowed nouns usually have one of these endings, e.g. Christian Alqosh *ʕamma* 'paternal uncle' (< Ar. *ʕamm*), *ʕašāya* 'dinner' (< Iraqi Ar. *ʕaša*) *ḥadāda* 'blacksmith' (< Ar. *ḥaddād*), *ʕāṣərta* 'early evening' (< Iraqi Ar. *ʕaṣir*) and *maʕwəlta* 'axe (or similar tool)' (< Iraqi Ar. *maʕwal* 'pickaxe'). Even if they do not, they are adapted to NENA stress patterns. Thus Ar. *ḥayawā́n* 'animal' is borrowed (possibly via N. Kurd. *ħeywan* [ħɛjˈwɑːn]) as *ḥɛwan* in Christian Alqosh, which has penultimate stress (Coghill 2004: 81).

More recent loans, on the other hand, may be used without any such modifications, e.g. Christian Alqosh *ʕamal* 'thing' (< Ar. *ʕamal* 'work'), *xām* 'linen' (Iraqi Ar. *ḫām* 'raw; cotton cloth'), and *sāʕa* 'hour' (f., < Ar. *sāʕa* f.). They often occur also in their original Arabic plural forms, e.g. Christian Alqosh *fallāḥín* 'farmers' and *ʔaʕdād* '(large) numbers'.

Many Arabic loanwords come with the Arabic feminine marker *tāʔ marbūṭa* (Standard Arabic *-a*). In *qəltu* Arabic dialects this usually has two realizations:

-a after emphatic or back consonants, otherwise a high vowel such as *-e* or *-i*.[13] Such loans in NENA usually also have the same distribution, that is *-e* (or the dialectal variant *-ə*), except after an emphatic or back consonant, when it is *-a* (Telkepe *-ɒ*), e.g. Christian Alqosh *baṭālə* 'idleness' and *rawḍa* 'kindergarten' and Christian Telkepe *ʕādə* 'custom' and *qəṣṣɒ* 'story' (see also §3.3.1).

Some loans appear to have come from Standard Arabic and have the *-a* regardless, e.g. Christian Telkepe *lahjɒ* 'dialect' and *madrasɒ* 'school'. Christian Qaraqosh seems to always represent the *tāʔ marbūṭa* as *-a* (Khan 2002: 204).

Borrowed nouns are quite commonly given Aramaic derivational suffixes. For instance, Jewish Azerbaijani *amona* 'paternal uncle' has a borrowed stem, *am-*, from Ar. *ʕamm* 'paternal uncle' via Kurdish or Azeri, but an Aramaic derivation, *-ona*, originally with diminutive function (Garbell 1965: 165). An example from the early Lišāna Deni texts is *ġaribūθa* 'foreignness', from Arabic *ɣarīb* 'foreign, strange' and the NENA abstract ending *-ūθa* (Sabar 1984: 205).

NENA often adopts the gender of the donor language, where that language has nominal genders (as in the case of Arabic and Northern Kurdish, which both have masculine–feminine gender systems). Thus, the following Christian Alqosh words share the same gender as their Arabic source: *ʕašāya* 'dinner' (m., like Iraqi Arabic *ʕaša*) and *daʔwa* 'wedding party' (f., like Arabic *daʕwa* 'invitation, party'). The loanword *ʕāṣərta* 'early evening' is, however, feminine (as indicated by the NENA feminine ending *-ta*), while the Arabic source (Iraqi Arabic *ʕaṣir*) is masculine. In Northern Kurdish, however, it is feminine (*'esir* [ʕæˈsɪr]), and this may have influenced the gender, which, in turn, motivated the adding of the feminine suffix.

In Christian Telkepe, some Arabic loanwords of the structure *CaCC have, when not suffixed, an epenthetic vowel between the final two consonants. This is absent when a suffix beginning with a vowel is added, i.e. the construct suffix *-ə̀d* or a possessive pronominal suffix. This follows the rules in the donor language: those Arabic dialects which have the epenthetic vowel (including Baghdadi and some *qəltu* dialects, such as Mosul) also lose it under similar conditions.[14] Examples include *ʕaqəl* 'mind': *ʕaql-əd=baxtɒ* [mind-CSTR=woman] 'a woman's mind'; and *ḥarub* 'war': *p-ḥarb-əd=sawāstipūl* [in-war-CSTR=Sebastopol] 'in the Crimean war'. It is interesting to note that the same rule is also found for Arabic loanwords in Kurdish (Thackston 2006: 5).

[13] See Jastrow (1979: 40) for the conditioned *imāla* (raising of a-vowels) in the *tāʔ marbūṭa* in the Arabic dialect of Mosul, and Jastrow (1990: 70) for the same in the Jewish Arabic dialect of ʕAqra and Arbīl.

[14] For Baghdadi Arabic, see Erwin (1963: 56–58).

Occasionally, loanwords are adapted to the native root-and-pattern templates, following the selection of a root. This frequently occurs when the root is also borrowed as a verb. Thus we find Christian Qaraqosh *ʔəjbona* 'a will, wish' (Khan 2002: 517), alongside the verb *√ʔjb* I 'to please' (< Ar. *√ʕǧb* IV), by analogy with native words on the pattern CəCCona, e.g. *yəqðona* 'a burn' (< *√yqð* I 'to burn'). Sabar (1984: 205) gives further examples from the early Lišāna Deni texts. More often, however, borrowed nouns are not adapted to native templates, e.g. Alqosh *ḥanafiya* 'tap' (< Ar. *ḥanafiyya*), or only coincidentally follow a native noun pattern (Arabic and NENA share many similar patterns), e.g. *qahwa* 'coffee' (< Ar. *qahwa*), which fits into the common Aramaic pattern CaCCa.

NENA dialects all have a variety of plural suffixes, the most common being perhaps *-e* (or its dialectal variant *-ə*). Loanwords, like inherited words, take a wide variety of native plural suffixes, but certain suffixes may be preferred or dispreferred for loanwords, in combination with other factors. For instance in Christian Alqosh feminine loanwords are not attested with the Aramaic plural suffixes *-wāθa* and *-awāθa*, while the loan-plural *-at* (< Ar. *-āt*) is almost exclusively found with loanwords (Coghill 2005: 347). Recent Arabic loans in Christian Nineveh Plain dialects often occur, unadapted, in their Arabic plural form (see §3.3.1).

3.1.3 Integration of adjectives

Like nouns, loan adjectives may occasionally be adapted to the native root-and-pattern templates, after the selection of a root. For instance, Arabic *ʔazraq* 'blue' (*√zrq*) is borrowed by Christian Alqosh as *zroqa* 'blue', by analogy with certain inherited colour adjectives of the form CCoCa, such as *smoqa* 'red'. Another example is Christian Alqosh *ʕadola* 'straight' (cf. Iraqi Ar. *ʕadil* 'straight' and Christian Qaraqosh which has borrowed it simply as *ʕadəl*).[15] More often the stem of the loan adjective is borrowed more or less unchanged, as in Christian Alqosh *faqira* 'poor' (Ar. *faqīr*), coincidentally fitting the inherited adjectival pattern CaCiCa. Adapted loan adjectives tend to take NENA inflection (e.g. f. *-ta~-θa*, pl. *-ə*). Unadapted loan adjectives usually take no inflection at all, e.g. Christian Telkepe *qə́rməzi* 'purple' (Ar. *qirmizī* m. 'crimson') and *ḍaʕíf* 'thin' (Iraqi Ar. *ḍaʕīf* m. 'thin, weak').

Loan-adjectives of a certain group including colours and bodily traits behave in a special manner in some NENA dialects: they take Aramaic inflection for

[15] Attested inherited words of the pattern CaCoCa are all in fact nouns in Christian Alqosh, e.g. *ʔalola* 'street'. The pattern CaCūCa might be more expected, being found with several common adjectives, e.g. *xamūṣa* 'sour'.

masculine and plural, but a special inflection *-ə* (identical to the plural ending) for the feminine. This occurs in Christian Qaraqosh particularly with Arabic loan adjectives, e.g. *ṭarša* 'deaf' (f./pl. *ṭaršə*, < Ar. m. *ʔaṭraš*, f. *ṭaršāʔ*) and *zarqa* 'blue' (f./pl. *zarqə*, < Ar. m. *ʔazraq*, f. *zarqāʔ*), see Khan (2002: 219). It appears to come from a dialectal reflex (*-ē*) of the Arabic *-āʔ* feminine ending, found especially with adjectives of these semantic groups.[16] In Christian Alqosh it is also found with loanwords of Northern Kurdish origin, e.g. *kačal-a* 'bald' (f./pl. *kačal-ə*, from N. Kurd. *k'eçel* [kʰæˈʧæl]).

In Arabic and Kurdish, adjectives normally follow the head noun, as in NENA. There are, however, a few pseudo-adjectival modifiers borrowed from Arabic which precede the noun in Arabic and are uninflected. These show the same behaviour when borrowed into NENA. One is *ʔawwal* 'first' in Christian Alqosh (a synonym to the inherited word *qamāya* 'first'), as in *ʔawwal=ga* 'the first time' – compare Arabic *ʔawwal marra* 'the first time'. Another is *ġer* 'other' (< Iraqi Ar. *ɣēr*), which is attested in Jewish Betanure, e.g. *ġer=məndi* 'something else' (Mutzafi 2008: 105) – compare Iraqi Arabic *ɣēr yōm* 'another day'. Another loanword, *xoš* 'good', invariably precedes the noun, e.g. Christian Telkepe *xoš=ʔixālɒ* 'good food'. This seems to originate in Iranian (Persian or Kurdish), but is also common in Iraqi and Anatolian Arabic dialects (as *ḫōš*), as well as in Turkic varieties (as *hoş* [hoʃ] or *xoş* [xoʃ]). In all these languages it precedes the noun, regardless of the usual word order.

3.1.4 Integration of verbs

The borrowing of verbs has been identified as potentially more complicated than the borrowing of other lexemes, due to their tendency to be morphologically complex (Matras 2009: 175). The borrowing of verbs in a Semitic language presents particular issues, due to the unusual root-and-pattern system. In Semitic languages verb lexemes are composed of a root (typically consisting of three – occasionally four – consonants or semi-vowels) and a derivation (also known as "stem", "form", "measure", "binyan" or "theme"). NENA dialects mostly have three triradical derivations (I, II and III) and at least one quadriradical derivation (Q). A borrowed verb will usually be integrated into this system. Three main strategies have been identified for the borrowing of verbs in NENA. One, common also in other Semitic languages (Wohlgemuth 2009: 173–180), is root extrac-

[16] Oddly enough, however, the realization as *-ē* seems to be restricted to Anatolian *qəltu* Arabic dialects (where it is stressed, e.g. Āzəḫ *lālé* 'dumb'), and not found in the dialects in Iraq (Jastrow 1978: 76). Other words ending in **-āʔ* in have *-ē* (unstressed) in *qəltu* Arabic dialects, but only as cases of *imāla* (raising of a-vowels) conditioned by a neighbouring high vowel.

tion, whereby from the phonological matter of the source verb a tri- or quadriradical root is selected. This is usually then allocated to a verbal derivation. A second is the borrowing of not only the root but also some of the morphology of the Arabic derivation: see below and §3.3.2. A third is the light verb strategy, whereby the loanverb consists of a light verb (with meanings such as 'become' or 'make') and a (verbal) noun, the latter containing the main semantic content.

The light verb strategy is found in some NENA dialects, but usually with Kurdish or Turkish verbs, which already consist of a light verb plus noun. It is not used to integrate Arabic loanverbs, although sometimes the noun in the predicate ultimately comes from Arabic.

The root-extraction strategy is well attested across NENA dialects and is particularly common with Arabic loanverbs. This is unsurprising, as these already have a root, which in many cases can simply be adopted as it is. For instance, Arabic *√ɣlb* I 'to win' (*ɣalaba* 'he won') is borrowed as Christian Telkepe *√ġlb* I 'to win'. Sometimes the root is adapted, to conform to the rules of root formation in NENA. For instance, 'geminate' roots, where the final two radicals are identical ($\sqrt{C_1C_2C_3}$, where C_2=C_3), are rare in NENA, and apparently absent altogether in derivation I. Just as inherited geminate roots were converted into middle-*y* roots ($\sqrt{C_1yC_3}$), so too are Arabic geminate roots. Thus, Arabic *√sdd* I 'to close, stop up' is borrowed as Christian Alqosh *√syd* I 'to close, seal' (compare inherited *√qyṛ* I 'to be cold' < *√qrr*).

Sometimes derivational affixes are adopted as radicals, often replacing a weak radical. For instance, Arabic derivation VIII verb *ittafaqa* (*√wfq*) is borrowed by Christian Alqosh as *√tfq* I 'to meet', with the VIII derivational infix *-t-* reanalysed as a radical. Frequently the root is borrowed not from a true verb but from a (verbal) noun or adjective. Thus, the NENA verb *√txmn* Q (found, e.g., in Jewish Betanure and Christian Qaraqosh, and as *√txml* Q in Alqosh) is borrowed from the Arabic noun *taḫmīn* (possibly via Northern Kurdish *t'exmîn* [tʰæxˈmiːn] 'supposition, guess'), itself a derivation of Arabic *√ḫmn* II 'to guess' (*ḫammana* 'he guessed'). The /t/ of the NENA root is not found in the Arabic root, but can only come from the verbal noun. This is an extension of an inherited Semitic strategy of deriving verbs from nouns. See Sabar (1984; 2002: 52) and Garbell (1965: 166) for more on the creation of verbal roots from non-Aramaic verbs.

The process of integration does not end with the establishment of a root, however. Every verb lexeme must also have a derivation. Tendencies can also be identified for this (Coghill 2015). Arabic loanverbs already have a derivation, but the majority of Arabic derivations have no cognate or functional equivalent in NENA. Where there is a cognate, there are also some formal and functional simi-

larities, and thus such cases are usually loaned into the cognate derivation. Thus, for instance, Arabic √ʕdl II (*ʕaddala*) 'to put in order' is borrowed as Christian Telkepe √ʃdl II 'to fix, tidy' (e.g. *mʃudəlli* 'I tidied'), Telkepe derivation II being the cognate of the Arabic derivation of the same number.

Verbs in Arabic derivations that have no cognate are sometimes allocated to derivations that bear some similarity in form or function to the original derivation. For instance, the NENA derivation most closely resembling Arabic derivation III in form is derivation II (the two share the template -CvCvC-, as opposed to -vCCvC-). Thus Arabic √*hǧr* III (*hāǧara*) 'to emigrate' is borrowed as Christian Telkepe √*hjr* II 'to emigrate' (e.g. *mhujera* 'they emigrated').

Arabic derivations VIII and X may be treated differently: in Christian Iraqi dialects, in particular those of the Nineveh Plain, the derivational morphology may itself be borrowed along with the lexeme (see §3.3.2).

3.1.5 Grammatical words and closed classes

NENA has freely borrowed grammatical words such as prepositions, conjunctions and particles of various functions, and some of these are Arabic, though most are Kurdish. In some cases, the original Arabic items may have been borrowed via Kurdish. In Christian Alqosh we find the preposition *ṣob* 'towards, near' (< Ar. *ṣawba* 'towards', cf. Iraqi Ar. *ṣōb* 'direction') and *baḥás* 'about, concerning' (< N. Kurd. *beḥs* [bæħs] 'discussion (about)' < Ar. *baḥθ*). Another example is *m-badal* 'instead of' (< *m-* 'from' + Iraqi Ar. *badāl*; Coghill 2004: 300). In Jewish Challa we also find *m-badal* and, in addition, *mābayn* 'between, among' (< Ar. *mā bayn*; Fassberg 2010: 149, 151). Even in Jewish Arbel, which generally shows less Arabic influence, we find *ḍidd* 'against' (< Ar. *ḍidd*; Khan 1999: 188).

Loan prepositions are not a new phenomenon in NENA, but are already attested in the early Jewish Lišāna Deni texts (Sabar 1984: 208), e.g. *ʕann-ıd* 'about' (< Ar. *ʕan* 'about'), *ṣōb* 'beside' (< Ar. *ṣawba*). By analogy with certain native prepositions, some have been extended with the construct suffix *-əd*, e.g. *ʕann-ıd*.

A particle that has been commonly borrowed is *bas* 'only; but' (cf. Iraqi Ar. *bass* 'enough; only; but'). This may have been borrowed via Northern Kurdish *bes* [bæs] 'enough; but'.

Many dialects, including Christian Alqosh and Christian Telkepe, use *kabira* to express 'much' or 'very'. This derives from Arabic *kabīr* 'big'. In Christian Qaraqosh (Khan 2002: 284–5) they use another Arabic loan for the same meaning: *ḥel ~ ḥelə* (cf. Iraqi Ar. *ḥēl* 'with force').

Other particles commonly borrowed are *fa* (roughly 'and so' in both Arabic and NENA) and *lo* 'or; either' (Iraqi Arabic *lō*). The adverb *baʕdén* 'then; later' (<

Ar. *baʃdēn*) is attested frequently in the Christian dialects of Alqosh, Telkepe and Qaraqosh, despite the presence of an inherited synonym, *baθər=dəx* [after=how] 'then; later'.

In Christian Alqosh and Christian Qaraqosh, a particle *də-* is used with imperatives to give the command a sense of urgency or encouragement. This is already attested in the early Jewish Lišāna Deni texts (Sabar 1976: xl). This appears to come from Northern Kurdish *de* [dæ] with the same function. A similar participle (*dē-, də-*) is found in both *qəltu* and Baghdadi Arabic (Jastrow 1978: 310–311).

3.2 Phonology

Two types of phonological contact influences in NENA will be considered here: new phonemes adopted through contact, and allophonic alternations influenced by contact.

3.2.1 New phonemes

NENA dialects have gained several new phonemes through language contact. These phonemes have entered the dialects via loanwords that were not fully adapted to Aramaic phonology. Some new phonemes are restricted to loanwords, while others have developed also in native words, through processes such as combination (creating affricate phonemes) and assimilation. As might be expected, Kurdish loanwords are responsible for the majority of the borrowed phonemes, but Arabic has also played a role, especially in those dialects closest to the Arabic-speaking region, i.e. the Christian dialects of the Nineveh Plain. The examples given below are from the Christian Alqosh dialect of this group (Coghill 2004: 11–25, with adapted transcription).

Some of the borrowed phonemes in NENA dialects have been introduced by both Kurdish and Arabic loanwords. These include /j/ [ʤ] and /č/ [ʧ]. The latter is not found in Standard Arabic, but is found in Mesopotamian dialects of Arabic. The phoneme /f/ seems to be borrowed predominantly from Arabic, although this phoneme also exists in Kurdish. Examples of loanwords with these three phonemes are: *ješ* 'army' (< Iraqi Ar. *ǧēš*), *jullə* 'clothes' (< N. Kurd. *cil* [ʤɪl]), *čārək* 'quarter' (< N. Kurd. *čarêk* [ʧɑːˈreːk]) *√čyk* I 'to pierce' (< Iraqi Ar. *√čkk* I), and *faqira* 'poor' (< Ar. *faqīr*).

The phoneme /č/ is also found in certain native Aramaic words, as a result of the combination of /t/ and /š/, e.g. *čeri* in *čeri qamāya* 'October' (< *tšeri, cognate with Christian Qaraqosh *təšri* and CSyr *tešri ~ tešrin* 'Tishrin').

The Arabic phoneme /ð̣/ [ðˤ] is found in many loanwords in Iraqi NENA dialects, e.g. *√ḥð̣r* III 'to prepare' (< Iraqi Ar. *√ḥð̣r* II). In most Mespotamian dialects of Arabic in contact with NENA, /ḍ/ is rarely found, as it has merged with /ð̣/. Nevertheless, one loanword in Alqosh and Qaraqosh has the /ḍ/ phoneme, namely *ʔoḍa* 'room', which originally comes from Turkish *oda*. While Turkish is not considered to have emphatic consonants, it does have vowel harmony, and words with back vowels have been interpreted as having emphatic consonants, when borrowed into *qəltu* (and other) Arabic dialects (Jastrow 1978: 51–52). Thus the *qəltu* dialect of Qarṭmin, in which *ḍ and *ð̣ have merged as /ð̣/, also has *ʔōḍa* 'room' (Jastrow 1978: 70). NENA *ʔoḍa* was borrowed from Turkish either via a local Arabic variety or directly, in which case its speakers must have also interpreted back-voweled Turkish words as emphatic.[17]

The pharyngeals /ʕ/ and /ḥ/, which in most inherited Aramaic lexemes have shifted to /ʔ/ and /x/ respectively, have been reintroduced through loanwords from both Arabic and the Classical Syriac used in the church. Examples for /ʕ/ are: *ʕamma* 'uncle' (< Ar. *ʕamm*), *√ʕyš* I 'to live' (< Ar. *√ʕyš* I), *ʕəddāna* 'time' (CSyr *ʕeddānā*). Examples for /ḥ/ are: *√jrḥ* I 'to get injured' (< Ar. *√ǧrḥ* I 'to injure'), *√ḥð̣r* III 'to prepare' (< Iraqi Ar. *√ḥð̣r* II), *mšiḥa* 'Christ' (< CSyr *mšiḥā*), and *ḥaṭṭāya* 'sinner' (< CSyr *ḥaṭṭāyā*). In some Arabic loans, however, /ʕ/ has shifted to /ʔ/, perhaps indicating that they belong to an earlier stratum, e.g. Christian Alqosh *daʔwa* 'wedding party' (Ar. *daʕwa*). Some cases of /ʕ/ and /ḥ/ in Alqosh, as in other NENA dialects, are original: the shift to /ʔ/ and /x/ respectively has been blocked in certain phonetic environments, particularly in the neighbourhood of emphatic consonants or /q/, e.g. *raḥūqa* 'far' (< *raḥḥūqa), see Khan (2002: 40–41). Furthermore, /ḥ/ has arisen in the third person singular possessive suffixes, as a shift from original *h. This appears to be a strategy of disambiguating these suffixes from the phonetically similar nominal endings (see Coghill 2008: 96–97).

The voiced uvular fricative was an allophone of the voiced velar stop /g/ in earlier Aramaic. In NENA it merged with *ʕ and shifted to a glottal stop /ʔ/. Like the pharyngeals, it has been reintroduced into NENA through loanwords from both Arabic and Classical Syriac, e.g. *√ġlb* I 'to win, defeat' (< Ar. *√ɣlb* I) and *paġra* 'body' (< CSyr *paḡrā*). It has also arisen in native words through regular assimilation of /x/ to a following voiced consonant. In the case of the verb *√ġẓd* I 'to reap' (< *√xẓd < *√xṣd < *√ḥṣd), the voiced allophone, originally only found in certain forms, has spread by analogy throughout the paradigm (Coghill 2004:

[17]Northern Kurdish also has this word, but Chyet's (2003) dictionary only gives variants without emphasis (e.g. *ode*), although Iraqi Kurdish dialects do often preserve emphasis in Arabic loanwords (Chyet 2003: viii; see also Öpengin, this volume).

20).

The cases of /č/, the pharyngeals, and /ġ/ show how new phonemes may arise through borrowing, while being assisted by internal developments.

3.2.2 Allophonic sound alterations

Some NENA dialects, such as Christian Alqosh (Coghill 2004: 27), exhibit final devoicing of voiced consonants, e.g. *mjāwəb* [mˈdʒæup] 'answer!' (cf. *mjawobə* 'to answer' with [b]) and *qapaġ* [ˈqɑpɐχ] 'lid' (cf. *qapaġəd-dəstiθa* 'saucepan lid', with [ʁ]). There is also a strong tendency towards devoicing in both *qəltu* Arabic (Jastrow 1978: 98) and the Kurdish dialects of Iraq (MacKenzie 1961: 49), so it seems to be an areal feature (see also Akkuş, this volume on contact-induced devoicing in Anatolian Arabic, and Lucas & Čéplö, this volume on the same phenomenon in Maltese).

3.3 Morphology

NENA dialects have borrowed a variety of morphemes from regional languages via lexical loans. As these become more integrated into the language, they may be found not only in the original loanwords but also with new words, including inherited lexemes. NENA being a Semitic language, it is possible for morphological borrowings to be a templatic pattern rather than a single phonetic chunk: indeed, some verbal derivational patterns have been borrowed from Arabic, as will be shown in §3.3.2.

3.3.1 Nominal inflection

A grammatical suffix that has been borrowed by some Iraqi dialects is the Arabic feminine sound plural suffix *-āt*. In Christian Alqosh and Christian Qaraqosh, as well as the Jewish Lišāna Deni dialects of northern Iraq, it has been integrated into the native morphology: as these dialects have penultimate stress in nouns, the suffix itself is not stressed in these dialects as it is in Arabic (Coghill 2004: 272–273; 2005; Khan 2002: 193–194). Accordingly it has also been shortened to *-at*, e.g. Christian Alqosh *makina* 'machine', pl. *makinat*, *maḥallə* 'town quarter', pl. *maḥallat*. In Alqosh and Qaraqosh it is only attested with feminine nouns. It is not, however, restricted to Arabic loans, but has been extended to other foreign words, e.g. Alqosh *pošiya* 'turban' (N. Kurd. *p'oşî* [pʰoːˈʃiː]) pl. *pošiyat*. In Alqosh and Qaraqosh it is even found with some native Aramaic words, e.g. Christian Qaraqosh *ʔarnuwa* 'rabbit', pl. *ʔarnuwat* 'rabbits'; *ʔilāna* 'tree', pl. *ʔilānat* 'trees'.

In some words, probably borrowed during the more recent and more intense period of contact with Arabic, the original stress and length of the ending is preserved, e.g. Christian Alqosh *holā́t* 'halls' and Christian Qaraqosh *badlā́t* 'suits' and *gadlā́t* 'tresses' (Khan 2002: 194). (Note, however, that the latter is an Aramaic word). This is always the case in Telkepe, e.g. *jəddɒ* 'midwife', pl. *jəddā́t* and *traktar* 'tractor', pl. *traktarā́t*. Note that in Telkepe, as in Arabic, this plural is sometimes found with masculine nouns, e.g. *mez* (m.) 'table', pl. *mezā́t* or *primuz* (m.) 'primus stove', pl. *primuzā́t*.

Apart from the Christian Nineveh Plain dialects, *-at* is attested regularly as a plural in some of the Jewish Lišāna Deni dialects, spoken further to the north. As mentioned in §2, these Jewish communities would have had contact with spoken Arabic through connections with their co-religionists.

In the modern Jewish dialect of Zakho, *-at* is used with the following types of nouns (Sabar 2002: 44–45): feminine Arabic loans ending in *-a* or *-e* (i.e. the dialectal version of the Arabic feminine suffix *tāʔ marbūṭa*; see §3.1.2), some nouns of Kurdish origin ending in *-e* (perhaps by analogy with Arabic loans ending in *-e*), and nouns ending in certain borrowed suffixes, namely the diminutive suffix *-ka* (f. *-ke*) borrowed from Kurdish, the professional suffix *-či* borrowed from Turkish, and the ending *-o*. It is also one of the two most common plurals for European loanwords, e.g. $^{+}$*pākētat* 'packets (of cigarettes)' (Sabar 1990: 57). This suggests it is particularly associated with loanwords, regardless of origin. In Jewish Duhok (also Lišāna Deni), however, it is attested with a native Aramaic word, *raʔolat* 'brooks' (Sabar 2002: 45). It seems therefore that the morpheme has been extended far beyond its original distribution.

The plural *-at* does not seem to have spread to all Lišāna Deni dialects, however: it is not mentioned in the grammars of Jewish Challa (Fassberg 2010) and Jewish Betanure (Mutzafi 2008). It has, nevertheless, an early origin: it is found in the late seventeenth-century manuscripts originating in the towns of ʕAmədya and Nerwa. I found one example of it in the grammar of the modern ʕAmədya dialect (Greenblatt 2011: 70), namely *maymonke* (f.) 'monkey', pl. *maymonkat*, probably because it has the Kurdish diminutive suffix (see above).

Across the border in Turkey, another Christian dialect has this plural ending, that is the dialect of ʕUmra (Turkish name *Dereköyü*), close to the town of Cizre. In this region of Turkey there are or were several Arabic-speaking communities, including Christian Arabic speakers in Cizre (until the First World War; see Jastrow 1978: 17), so it is not surprising that there should be influence from Arabic. In this dialect, *-at* is mostly attested with borrowed feminine nouns ending in *-e*, though there are also a couple ending in *-a*, both masculine and feminine

(Hobrack 2000: 114). The majority have the Kurdish diminutive suffix *-ka* (f. *-ke*) mentioned above in relation to Jewish Zakho.

In the Christian dialects of Iraq, as spoken currently, it is common to use Arabic words with their original plural morphology, probably because almost all speakers speak Arabic with native or near-native competence and many concepts are more familiar or only available to them in this language.[18] Thus, apart from the *-āt* plural, we also find the masculine sound plural suffix *-in* and the non-concatenative broken plurals, e.g. Christian Alqosh *fallāḥ-ín* 'farmers', and *barāmíl* 'barrels' (sg. *barmíl*) (Coghill 2004: 273). We even find such examples in the late seventeenth-century manuscripts written in Jewish Lišāna Deni dialects, e.g. *ġāfilīn* 'fools' and *ʔarwāḥ* 'spirits' (Sabar 1984: 205–206).

Many Arabic loanwords come with the Arabic feminine marker *tāʔ marbūṭa*, either the *qəltu* Arabic variants or the Standard Arabic *-a* (§3.1.2). In some dialects of the Nineveh Plain, the *tāʔ marbūṭa* is borrowed along with its connecting allomorph *-ət*. In Arabic the /t/ is only realized in construct state (as the head of a genitive phrase) or before possessive suffixes.

In Christian Qaraqosh the isolated form of such loans ends in *-a*, like inherited masculine nouns, although the gender is feminine (as in the source words). When possessive suffixes are added, however, the /t/ is realized, as in Arabic (Khan 2002: 204–206). Thus Qaraqosh *badla* 'suit of clothes' (cf. Iraqi Arabic *badla*) becomes *badl-ətt-əḥ* [suit-F-3SG.M] 'his suit of clothes'. The gemination of the /t/ is not found in the Arabic forms, but can be explained as follows. In Mosul Arabic, unlike in many Arabic dialects, the *tāʔ marbūṭa* takes the stress, when any possessive suffix is added: *báṣali* 'onion', *baṣal-át-ak* [onion-F-2SG.M] 'your onion' (Jastrow 1983: 105). It is likely that the /ə/ vowel in the NENA morpheme *-ətt-* imitates the vowel of the Arabic morpheme. The stress pattern fits well into NENA, which has penultimate stress. However, in NENA /ə/ is dispreferred in an open syllable, especially when stressed. The /t/ is probably geminated in order to close the syllable so as to conform to this preference.[19] This mechanism has parallels elsewhere in NENA.

These same loanwords take the Arabic plural *-at* discussed above. Even some Aramaic feminine words in Christian Qaraqosh have acquired both *-ətt-* and *-at*, e.g. *ʔarnuwa* (f.) 'rabbit', *ʔarnuwəttəḥ* 'his rabbit', *ʔarnuwat* 'rabbits'. But *-ətt-* is also found with some Aramaic feminine words that have native plurals, e.g. *bira*

[18]Younger NENA speakers who have grown up in the Kurdish-controlled region since 1991 may have less competence in Arabic, however.

[19]Khan (2002: 206) gives two other possible derivations: a combination of Arabic f. *-ət* and Aramaic f. *-ta* (though the latter is not found on the isolated form) or the NENA independent genitive particle *did-*. The explanation above seems to me to be simpler, however.

(f.) 'well', *birāθa* 'wells', *birəttəḥ* 'his well'. In exceptional cases *-ətt-* may also be used with feminine words with the Aramaic f. ending *-ta~-θa*, e.g. *šwiθa* 'bed', *šwiyāθa* 'beds', *šwiθəttəḥ* 'his bed'. It seems, therefore, that in Qaraqosh this is now a morphological borrowing independent of the loanwords it was originally borrowed with.

In Christian Telkepe, vernacular Arabic nouns with *tāʔ marbūṭa* are borrowed ending in either *-ə* or *-ɒ*, matching the two realizations of the *tāʔ marbūṭa* in *qəltu* Arabic (§3.1.2). As in Qaraqosh, these nouns retain their feminine gender in Telkepe. They also have the *-ətt-* allomorph before possessive suffixes, e.g. *ṣəḥḥɒ* (f.) 'health', *ṣəḥəttux* [*ṣəḥ-ətt-ux* health-F-2SG.M] 'your (m.) health'; *qubbə* (f.) 'room', *qubbətte* [*qubb-ətt-e* room-F-3SG.M] 'his room'. The infix seems to be used productively with Arabic words, as and when they are used. One example in Telkepe is not borrowed from a feminine with *tāʔ marbūṭa*, namely *čāyi* (f.) 'tea' (cf. Iraqi Ar. *čāy* (m.)). This word is, however, feminine in Northern Kurdish (*çay* [ʧɑːj]), whence it may have been borrowed.

Christian Alqosh seems to have gone a step further, creating back-formations from the suffixed forms. Thus the unsuffixed forms also have *-ətt-*, e.g. *ṣaḥətta* 'health', *qaṣətta* 'story' and *məllətta* 'religious community'. When the plural suffix (always the feminine plural *-yāθa*) is added, one /t/ alone is preserved, suggesting that the second is now analysed as part of the feminine singular ending *-ta*, while *-ət-* is analysed as part of the stem*:* *qaṣət-ta* 'story', *qaṣət-yāθa* 'stories'; *məllət-ta* 'community', *məllət-yāθa* 'communities'.

Similar forms are also attested in Jewish Challa (Lišāna Deni), but without the gemination of the /t/, e.g. *məlləta* 'ethnic group', *ʕādəta* 'custom' (Fassberg 2010: 52). Rather than explaining the /t/ as originating in the Arabic suffixed stem, as I have done above, Fassberg suggests that the /t/ is present because the words were borrowed via (Northern) Kurdish, which realizes the *tāʔ marbūṭa* as a final /t/ even when the noun is unsuffixed: *milet* [mɪˈlæt] and *ʕadet* [ʕɑːˈdæt] (Chyet 2003: 387). Khan (2002: 206) also suggests this route for Qaraqosh. This explanation would not explain why the unaffixed forms in Qaraqosh do not end in /t/, nor why the preceding vowel in all these dialects is /ə/ rather than /a/ (the nearest phonetic equivalent to Kurdish ⟨e⟩). In fact, there are some clear loans of Arabic words via Kurdish which end in *-at* in the singular unsuffixed form (see §3.1.1). The Kurdish route would furthermore not explain the close association in Qaraqosh of this morpheme with words taking an *-at* plural, which seems to have been borrowed directly from Arabic. It seems more likely, therefore, that the Qaraqosh, Alqosh and Challa feminine nouns with infixed *-ət-* have been borrowed directly from Arabic and are influenced by the Arabic suffixed forms,

which have this infix.

3.3.2 Verbal derivation

The NENA verbal system consists of both synthetic and analytic verb forms. The synthetic verb forms are formed from two stems, the Present Base and the Past Base, e.g. Christian Alqosh *k-šaql-i* [IND-take.PRES-3PL] 'they take' and *šqəl-lɛ* [take.PAST-3PL] 'they took'. Analytic forms involve auxiliary verbs or verboids combined with non-finite verb forms, such as the infinitive or participles, or, less often, with finite verb forms. Like Arabic, NENA has a verbal system based on the root-and-pattern system. As also in Arabic, a verb lexeme typically has a triconsonantal root and a verbal derivational class (see §3.1.4). While Standard Arabic has ten fairly common triradical verbal derivations, NENA dialects typically have only three or four inherited verbal derivations.

Morphological loans may be found in the verbal system. Christian NENA dialects of the Nineveh Plain and elsewhere have partially borrowed Arabic verbal derivations along with borrowed verb lexemes. NENA and Arabic have some cognate verbal derivations and the relationships are relatively transparent. Most Arabic loanverbs are allocated to a NENA derivation that is formally or functionally similar to the donor derivation (and often cognate). See §3.1.4 for discussion of this. In the case of Arabic verbal derivations VIII and X, however, this is not possible, as no NENA derivations have the characteristic affixes *-t-* and *(i)st-*. In some cases, the affix may instead be analysed as a radical (§3.1.4). In others, loanverbs in these derivations are borrowed with this derivational morphology, i.e. with the affixes. This has, in effect, created new derivations, the Ct- and St-derivations.

Table 1 gives all hitherto attested examples of verbs in the new derivations from Christian Telkepe, but additional verbs are attested in Christian Qaraqosh (Khan 2002: 130).

When Arabic verbs in derivations VIII and X are borrowed as they are, their characteristic consonantal clusters *-Ct-* and *-st-* are preserved and not broken up by an epenthetic vowel, even if this results in a syllabic structure that is dispreferred in the NENA dialect (such as a stressed short vowel in an open syllable), e.g. *k-maḥtarəm* [IND-respect.PRES.3SG.M] 'he respects'. This may be in order to preserve a salient characteristic of the original Arabic forms.

The vowel pattern in these derivations is, on the other hand, variable, even within the speech of one speaker. For instance, in the Present Base of the St-derivation, we find məstaCaCC-, məstaCCəC- and məstaCəCC- (e.g. *məstaʃaml-, məstaʃməl-, məstaʃəml-* 'use') as variants of one and the same form. What are

Table 1: Arabic loanverbs borrowed into the new NENA derivations

NENA verb	Source verb
√*ḥrm* Ct- 'to respect'	Ar. √*ḥrm* VIII (*iḥtarama*)
√*xlf* Ct- 'to differ'	Ar. √*ḫlf* VIII (*iḫtalafa*)
√*ḥfl* Ct- 'to celebrate'	Ar. √*ḥfl* VIII (*iḥtafala*)
√*ʕml* St- 'to use'	Ar. √*ʕml* X (*istaʕmala*)
√*ġll* St- 'to exploit'	Ar. √*ɣll* X (*istaɣalla*)

the reasons for this variability? Firstly, Arabic derivations VIII and X are morphophonemically more complex than the native Aramaic derivations. The consonant clusters bring the necessity of epenthetic vowels: this leads to at least one short vowel in an open syllable, which is disfavoured in Telkepe. Where the epenthetic vowel is placed is still optional and in flux. Secondly, there is a conflict between the characteristic vowels of the Iraqi Arabic source and the vowels typical of Aramaic derivations. Sometimes the former may be more influential and sometimes the latter.

The new Ct- and St- derivations in NENA have not been extended to inherited roots nor used productively, unlike some Arabic derivations in Western Neo-Aramaic. See Coghill (2015) for full details of the new derivations found in NENA, Western Neo-Aramaic and other Neo-Aramaic varieties.

3.4 Syntax and pattern borrowings

A syntactic borrowing attested only in the Christian Nineveh Plain dialects is the grammaticalization of a prospective auxiliary (and, as a further step, uninflected particle) on the model of the vernacular Arabic prospective future particle *raḥ-*, which is attested in nearby Mosul Arabic (author's fieldwork), as well as more widely across the Syrian and Mesopotamian Arabic dialects (Jastrow 1978: 304). Example (1) shows the Neo-Aramaic construction (with the particle) and example (2) shows the Arabic construction.

(1) Christian Telkepe NENA (author's fieldwork)
zi-napl-ɒ
PRSP-fall.PRES-3SG.F
'She's going to fall.'

(2) Christian Mosul Arabic (author's fieldwork)
ɣāḥ-təqaʕ š-šağaɣa!
PRSP-fall.IMPF.3SG.F DEF-tree
'The tree's going to fall!'

In both cases the gram has developed from a verb 'to go' in a form with imperfective or imperfective-like functions.[20] Such a development is of course extremely common in the world's languages and does not need a contact explanation. Nevertheless, there is evidence that contact played a role. The construction is only found in NENA dialects close to the Arabic-speaking zone of Iraq, i.e. near to Mosul. Furthermore, the most mature versions of the gram (formally and functionally) are found in the villages closest to Mosul. The gram seems to have developed only in the last 100 years or so, as it is not attested in texts or mentioned in grammars of those dialects before then. See Coghill (2010; 2012) for more details.

NENA shares a number of idiomatic expressions with neighbouring languages. Among these are formulae used regularly in specific contexts, such as telling a story or expressing thanks, congratulations or condolences. One that is widespread in NENA dialects, as well as several neighbouring languages, is the opening formula to a fictional story, which begins 'there was (and) there wasn't': see also Chyet (1995: 236–237). It is attested in various dialects of NENA, Ṭuroyo, Kurdish, Azeri, Persian and Arabic, e.g.:

Christian Alqosh NENA *ʔəθwa=w laθwa* (Coghill 2009: 268)
Christian Bohtan NENA *ətwa lətwa* (Fox 2009)
Akre Kurdish *hebo nebo* [hæˈbo: næˈbo:] (MacKenzie 1962: 288)
Iranian Azeri *(bir) vármiš (bir) jóxmuš* (Garbell 1965: 175)
Christian Bəḥzāni Arabic *kān w ma kān* (Jastrow 1981: 404)[21]

When such formulae are shared by multiple regional languages, it is difficult to say for certain which language NENA borrowed them from. Kurdish is usu-

[20] In the case of the Nineveh Plain dialects, it originates in a verb that originally had perfect aspect, e.g. *zil-ən* 'I have gone', possibly with the implication of 'I am on my way'. It had also acquired a meaning of imminent future 'I am about to go', in effect 'I am in the process of just leaving', hence "imperfective-like functions".

[21] This is a variant (along with *kān ma kān*, attested in Palestinian Arabic) of the well-known formula *kān yā ma kān* 'once upon a time'. While *kān* ***w*** *ma kān* clearly means 'there was and there was not', *kān* ***yā*** *ma kān* has been interpreted in different ways both by scholars and native speakers. Taking *yā ma* in its meaning of 'how much', it can be understood as 'there was, how much there was!' Alternatively, the *ma* is understood as a negator, as is found in the formula in the other languages. See Lentin (1995) for a discussion of *kān yā ma kān* and similar expressions.

ally the assumed donor, simply because it is the language most in contact with NENA and which has had the greatest influence at all levels. Given, however, that many speakers knew other regional languages as well, they may have heard such expressions in several languages.

Proverbs are another area in which there are shared expressions across the regional languages (Chyet 1995: 234–236; Garbell 1965: 175; Segal 1955). An example is 'He who knows, knows. He who doesn't know, says "a handful of lentils".' This stems from a folktale and means something like 'looks can be deceiving' (Chyet 1995: 235–236). It is attested in Kurdish, Iraqi Arabic, and NENA, as illustrated in (3)–(4).

(3) Iraqi Arabic (Chyet 1995: 235)
il-yidrī yidrī w-il ma yidrī
REL-know.IMPF.3SG.M know.IMPF.3SG.M and-REL NEG know.IMPF.3SG.M
gaḏbit ʕadas
handful.CS lentils
'He who knows knows, he who doesn't know (says) "a handful of lentils".'

(4) Jewish Zakho NENA (Segal 1955: 262, adapted transcription)
aw d-k-īʔe k-īʔe aw d-lá
3SG.M REL-IND-know.PRES.3SG.M IND-know.PRES.3SG.M 3SG.M REL-not
k-īʔe g-mēnüx bi-ṭloxe
IND-know.PRES.3SG.M IND-look.PRES.3SG.M at-lentils
'He who knows knows, he who doesn't know looks at a handful of lentils.'

Sabar (1978), who lists proverbs used by the Jews of Zakho, states also that many proverbs were not translated into NENA, but used in the original language, whether Kurdish or Arabic.

There are also some areas of structural convergence in the region's languages, where the donor language cannot be definitely identified. For instance, all the languages (NENA, Sorani, Northern Kurdish, Persian, Turkish, Azeri, Iraqi Turkman and *qəltu* Arabic) have enclitic copulas, as illustrated in (5)–(7).

(5) Akre Kurdish (MacKenzie 1961: 175)
ew kî=e [æw ˈkiːæ]
DEM who=PRS.COP.3SG
'Who is that?'

(6) Christian Telkepe NENA (author's fieldwork)
man=ilə
who=PRS.COP.3SG.M
'Who is he?'

(7) Jewish Arbel Arabic (Jastrow 1990: 37, 46)
mani=we
who=3SG.M
'Who is he?'

Another shared structure is the use of finite subordinate clauses in subjunctive mood, rather than infinitives, as complements. In earlier Aramaic varieties, such as Classical Syriac, both were used (Nöldeke 1904: 224–226), but in NENA only finite verbs are used, as in example (8).

(8) Christian Telkepe NENA
k-əbə d-āxəl
IND-want.PRES.3SG.M COMP-eat.PRES.3SG.M
'He wants to eat.'

Finite verbs in an irrealis mood are also used in such subordinate clauses in *qəltu* (and other vernacular) Arabic (e.g. Jastrow 1990: 65), Northern Kurdish (MacKenzie 1961: 208–209), Sorani (MacKenzie 1961: 134–135), Iraqi Turkman (Bulut 2007: 175–176), and Iranian Azeri (Fariba Zamani, personal communication). The development in Turkic is attributed to Iranian influence (Bulut 2007: 175–176). This parallels the loss of the infinitive and its replacement by finite verb forms in the Balkan Sprachbund (see, e.g., Joseph 2009).

The existence of markers in the noun phrase to specify for indefiniteness (and in many cases specificity, e.g. 'a certain man') is widespread in the area, being found in NENA (*xa-* 'one, a (certain)'), Northern Kurdish (*-ek* [ɛk] < *yek* 'one'), Sorani (*-ēk* [e:k] < *yek*), *qəltu* Arabic (*faɣəd* < *fard* 'individual'), Baghdadi Arabic (*fadd* < *fard*) and Turkish/Azeri (*bir* 'one').

4 Conclusion

Though not the dominant contact language, Arabic has influenced NENA dialects considerably, especially those in close contact with Arabic-speaking population centres, namely in the Christian Nineveh Plain dialects, the Jewish Lišāna Deni dialects and in the Christian dialects in Şırnak province in Turkey.

The influence from Arabic is manifested mostly in lexicon, phonology and morphology, and less in syntax.

Arabic influence has occurred in different phases. Earlier Arabic influence was mostly indirect, via Kurdish loans, but direct borrowing seems to have occurred too.

In the twentieth and twenty-first centuries, Arabic influence has increased dramatically in the dialects spoken in Iraq, due to mass education exclusively in Arabic, as well as national media, military service, improved transport, and migration to the Iraqi cities. Most NENA speakers are bilingual and speak Arabic with native competence, and this has affected how they use Arabic words within their own language. Typically, recent loans are unadapted and close to code-switching.

As much of the fieldwork on which this description depends was undertaken in the late twentieth century or first few years of the twenty-first century, in future research it would be interesting to look at the speech of young people today and see whether much has changed. It would also be worth comparing the speech of communities in their ancestral villages with diaspora communities living in (or who have recently left) Baghdad or Basra.

Further reading

Most work on NENA and language contact has focused on contact with Kurdish. To my knowledge, only three works are dedicated to contact with Arabic, none of which is an overview: Sabar's (1984) study of Arabic influence in the early texts in Jewish Lišāna Deni; Coghill's (2010; 2012) research into a prospective construction found in the Christian Nineveh Plain dialects, which has apparently grammaticalized under influence from Arabic; and Coghill's (2015) study of new verbal derivations borrowed from Arabic into various Neo-Aramaic languages, including NENA.

Khan's (2002) grammar of Christian Qaraqosh contains a great deal of information, scattered through the volume, about contact influences from Arabic, Qaraqosh being one of the dialects most affected by such influence.

Acknowledgements

I would like to thank all the Neo-Aramaic speakers who have generously given of their time in my fieldwork and the fieldwork of other scholars which is cited in this chapter. Much of my research on NENA and language contact took place

in the German Research Foundation-funded project, *Neo-Aramaic morphosyntax in its areal-linguistic context.* I would also like to thank the editors of this volume for their helpful comments.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	M/m.	masculine
Ar.	Arabic	N. Kurd.	Northern Kurdish
COMP	complementizer	NEG	negator
COP	copula	PAST	NENA Past Base
CS	construct state	PL/pl.	plural
CSyr	Classical Syriac	PRES	NENA Present Base
DEM	demonstrative	PRS	present
F/f.	feminine	PRSP	prospective
IMPF	Imperfect (prefix conjugation)	REL	relativizer
IND	indicative	SG	singular

Symbols

I, II, III etc.	Arabic verbal derivations
I, II, III, Q	NENA verbal derivations
=	links two words or morphemes in a phrase with a single stress on the second component (including but not limited to proclitics)
=	links two words or morphemes in a phrase with a single stress on the first component (including but not limited to enclitics)

References

Blanc, Haim. 1964. *Communal dialects in Baghdad.* Cambridge, MA: Harvard University Press.

Brauer, Erich & Raphael Patai. 1993. *The Jews of Kurdistan.* Detroit: Wayne State University Press.

Bulut, Christiane. 2007. Iraqi Turkman. In J. Nicholas Postgate (ed.), *Languages of Iraq, ancient and modern*, 159–187. London: British School of Archaeology in Iraq.

Chyet, Michael L. 1995. Neo-Aramaic and Kurdish: An interdisciplinary consideration of their influence on each other. In Shlomo Izre'el & Rina Drory (eds.), *Language and culture in the Near East*, 219–252. Leiden: Brill.

Chyet, Michael L. 2003. *Kurdish–English dictionary*. New Haven, CT: Yale University Press.

Coghill, Eleanor. 2004. *The Neo-Aramaic dialect of Alqosh*. Cambridge: University of Cambridge. (Doctoral dissertation).

Coghill, Eleanor. 2005. The morphology and distribution of noun plurals in the Neo-Aramaic dialect of Alqosh. In Alessandro Mengozzi (ed.), *Studi afroasiatici: XI Incontro italiano di linguistica camitosemitica*, 337–348. Milan: Franco-Angeli.

Coghill, Eleanor. 2008. Some notable features in North-Eastern Neo-Aramaic dialects of Iraq. In Geoffrey Khan (ed.), *Neo-Aramaic dialect studies: Proceedings of a workshop on Neo-Aramaic held in Cambridge, 2005*, 91–104. Piscataway, NJ: Gorgias Press.

Coghill, Eleanor. 2009. Four versions of a Neo-Aramaic children's story. *ARAM Periodical* 21. 251–280.

Coghill, Eleanor. 2010. The grammaticalization of prospective aspect in a group of Neo-Aramaic dialects. *Diachronica* 27(3). 359–410.

Coghill, Eleanor. 2012. Parallels in the grammaticalisation of Neo-aramaic *zil-* and Arabic *raḥ-* and a possible contact scenario. In Domenyk Eades (ed.), *Grammaticalization in Semitic*, 127–144. Oxford: Oxford University Press.

Coghill, Eleanor. 2014. Differential object marking in Neo-Aramaic. *Linguistics* 52(2). 335–364.

Coghill, Eleanor. 2015. Borrowing of verbal derivational morphology between Semitic languages: The case of Arabic verb derivations in Neo-Aramaic. In Francesco Gardani, Peter Arkadiev & Nino Amiridze (eds.), *Borrowed morphology*, 83–108. Berlin: De Gruyter Mouton.

Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic*. Washington, DC: Georgetown University Press.

Fassberg, Steven Ellis. 2010. *The Jewish Neo-Aramaic dialect of Challa*. Leiden: Brill.

Fox, Samuel Ethan. 2009. *The Neo-Aramaic dialect of Bohtan*. Piscataway, NJ: Gorgias Press.

Garbell, Irene. 1965. The impact of Kurdish and Turkish on the Jewish Neo-Aramaic dialect of Persian Azerbaijan and the adjoining regions. *Journal of the American Oriental Society* 85(2). 159–177.

Greenblatt, Jared R. 2011. *The Jewish Neo-Aramaic dialect of Amədya*. Leiden: Brill.

Hoberman, Robert D. 1989. *The syntax and semantics of verb morphology in modern Aramaic: A Jewish dialect of Iraqi Kurdistan*. New Haven, CT: American Oriental Society.

Hobrack, Sebastian. 2000. *Der neuaramäische Dialekt von Umra (Dereköyü): Laut- und Formenlehre – Texte – Glossar*. Erlangen–Nürnberg: Friedrich-Alexander-Universität. (Magister dissertation).

Jastrow, Otto. 1978. *Die mesopotamisch-arabischen* qəltu*-Dialekte*. Vol. 1: *Phonologie und Morphologie*. Wiesbaden: Steiner.

Jastrow, Otto. 1979. Zur arabischen Mundart von Mossul. *Zeitschrift für Arabische Linguistik* 2. 76–99.

Jastrow, Otto. 1981. *Die mesopotamisch-arabischen* qəltu*-Dialekte*. Vol. 2: *Volkskundliche Texte in elf Dialekten*. Wiesbaden: Steiner.

Jastrow, Otto. 1983. Tikrit Arabic verb morphology in a comparative perspective. *Al-Abhath* 31. 99–110.

Jastrow, Otto. 1990. *Der arabische Dialekt der Juden von ʿAqra and Arbīl*. Wiesbaden: Harrassowitz.

Joseph, Brian D. 2009. *The synchrony and diachrony of the Balkan infinitive: A study in areal, general and historical linguistics*. Cambridge: Cambridge University Press.

Khan, Geoffrey. 1999. *A grammar of Neo-Aramaic: The dialect of the Jews of Arbel*. Leiden: Brill.

Khan, Geoffrey. 2002. *The Neo-Aramaic dialect of Qaraqosh*. Leiden: Brill.

Khan, Geoffrey. 2007. Grammatical borrowing in North-eastern Neo-Aramaic. In Yaron Matras & Janet Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 197–214. Berlin: Mouton de Gruyter.

Khan, Geoffrey. 2008. *The Neo-Aramaic dialect of Barwar*. Leiden: Brill.

Khan, Geoffrey. 2009. *The Jewish Neo-Aramaic dialect of Sanandaj*. Piscataway, NJ: Gorgias Press.

Lentin, Jérôme. 1995. *kān ya ma kān*: sur quelques emplois de *ma* dans les dialectes arabes du moyen-orient. *Studia Orientalia Electronica* 75. 151–162.

MacKenzie, David N. 1961. *Kurdish dialect studies*. Vol. 1. Oxford: Oxford University Press.

MacKenzie, David N. 1962. *Kurdish dialect studies*. Vol. 2. Oxford: Oxford University Press.

Maclean, Arthur John. 1895. *Grammar of the dialects of vernacular Syriac as spoken by the eastern Syrians of Kurdistan, north-west Persia, and the Plain of Mosul: With notices of the vernacular of the Jews of Azerbaijan and of Zakhu near Mosul*. Cambridge: Cambridge University Press.

Maclean, Arthur John. 1901. *A dictionary of the dialects of vernacular Syriac as spoken by the Eastern Syrians of Kurdistan, northwest Persia, and the plain of Moṣul*. Oxford: Clarendon press.

Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.

Mengozzi, Alessandro. 2002. *"Israel of Alqosh and Joseph of Telkepe: A Story in a Truthful Language": Religious poems in vernacular Syriac (North Iraq, 17th century)*. Vol. 2: *Introduction and translation*. Leuven: Peeters.

Mutzafi, Hezy. 2004. *The Jewish Neo-Aramaic dialect of Koy Sanjaq (Iraqi Kurdistan)*. Wiesbaden: Harrassowitz.

Mutzafi, Hezy. 2008. *The Jewish Neo-Aramaic Dialect of Betanure (Province of Dihok)*. Wiesbaden: Harrassowitz.

Nöldeke, Theodor. 1904. *Compendious Syriac grammar*. London: Williams & Norgate.

Rizgar, Baran. 1993. *Dictionary Ferheng Kurdish–English English–Kurdish (kurmancî)*. London: M. F. Onen.

Sabar, Yona. 1976. pəšaṭ wayəhî bəšallaḥ: *a Neo-Aramaic midrash on* Beshallaḥ *(Exodus): Introduction, phonetic transcription, translation, notes, and glossary*. Wiesbaden: Harrasowitz.

Sabar, Yona. 1978. Multilingual proverbs in the Neo-Aramaic speech of the Jews of Zakho, Iraqi Kurdistan. *International Journal of Middle East Studies* 9(2). 215–235.

Sabar, Yona. 1984. The Arabic elements in the Jewish Neo-Aramaic texts of Nerwa and ʿAmādīya, Iraqi Kurdistan. *Journal of the American Oriental Society* 104(1). 201–211.

Sabar, Yona. 1990. General European loanwords in the Jewish Neo-Aramaic dialect of Zakho, Iraqi Kurdistan. In Wolfhart Heinrichs (ed.), *Studies in Neo-Aramaic*, 53–66. Atlanta, GA: Scholars Press.

Sabar, Yona. 2002. *A Jewish Neo-Aramaic dictionary: Dialects of Amidya, Dihok, Nerwa and Zakho, northwestern Iraq*. Wiesbaden: Harrassowitz.

Segal, Judah B. 1955. Neo-Aramaic proverbs of the Jews of Zakho. *Journal of Near Eastern Studies* 14(4). 251–270.

Sinha, Jasmin. 2000. *Der neuostaramäische Dialekt von Bēṣpən (Provinz Mardin, Südosttürkei): Eine grammatische Darstellung*. Harrassowitz.

Talay, Shabo. 2011. Arabic dialects of Mesopotamia. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 909–919. Berlin: De Gruyter Mouton.

Thackston, Wheeler M. 2006. *Kurmanji Kurdish: A reference grammar with selected readings*. Harvard University.

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.

Wohlgemuth, Jan. 2009. *A typology of verbal borrowings*. Berlin: Mouton de Gruyter.

Woodhead, Daniel R & Wayne Beene. 1967. *A dictionary of Iraqi Arabic: Arabic–English*. Washington, DC: Georgetown University Press.

Chapter 7

Iranian languages

Dénes Gazsi

Iranian languages, spoken from Turkey to Chinese Turkestan, have been in language contact with Arabic since pre-Islamic times. Arabic as a source language has provided phonological and morphological elements, as well as a plethora of lexical items, to numerous Iranian languages under recipient-language agentivity. New Persian, the most significant member of this group, has been a prominent recipient of Arabic language elements. This study provides an overview of the historical development of this contact, before analyzing Arabic elements in New Persian and other New Iranian languages. It also discusses how Arabic has influenced Modern Persian dialects, and how Persian vernaculars in the Persian Gulf region of Iran have incorporated Arabic lexemes from Gulf Arabic dialects.

1 Current state and historical development

1.1 Iranian languages

Iranian languages, along with Indo-Aryan and Nuristani languages, constitute the group of Indo-Iranian languages, which is a sizeable branch of the Indo-European language family. The term "Iranian language" has historically been applied to any language that descended from a proto-Iranian parent language spoken in Central Asia in the late third to early second millennium BCE (Skjærvø 2012).

Iranian languages are known from three chronological stages: Old, Middle, and New Iranian. Persian is the only language attested in all three historical stages. New Persian, originally spoken in Fārs province, descended from Middle Persian, the language of the Sasanian empire (third–seventh centuries CE), which is the progeny of Old Persian, the language of the Achaemenid empire (sixth–fourth centuries BCE). New Persian is divided into Early Classical (ninth–twelfth centuries CE), Classical (thirteenth–nineteenth centuries) and Modern Persian

Dénes Gazsi. 2020. Iranian languages. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 157–174. Berlin: Language Science Press. DOI:??

(from the nineteenth century onward), the latter considered to be based on the dialect of Tehran (Jeremiás 2004: 427).

Today, Iranian languages are spoken from the Caucasus, Turkey and Iraq in the west to Pakistan and Chinese Turkestan in the east, as well as in a large diaspora in Europe and the Americas. New Iranian languages are divided into two main groups: Western and Eastern Iranian languages. The focus of this study is New Persian, the most significant member among Iranian languages, but a brief overview of Arabic influence on other New Iranian languages will also be provided. Below is a list of the most important members and their geographical distribution (Schmitt 1989: 246).

1.1.1 Western Iranian languages

1.1.1.1 Southwestern group

Persian (*Fārsī*) (spoken throughout Iran and adjacent areas), Tajik (the variety of New Persian in Central Asia), Darī Persian (Afghanistan), Kumzārī (Musandam Peninsula). Persian dialects in this group include Dizfūlī (Khuzestan province), Lurī (ethnic group along the Zagros mountain range), Baḫtiārī (nomadic tribe in the Zagros mountains), Fārs dialects (Fārs province), Lāristānī dialects (Lāristān region of Fārs province), Bandarī (dialects spoken around Bandar ʿAbbās in the Persian Gulf region, to which Fīnī also belongs).

1.1.1.2 Northwestern group

Kurdish, Zazaki (in eastern Turkey), Gurānī (in eastern Iraq and western Iran), Balūčī (Balochi, spoken chiefly in Iranian and Pakistani Baluchistan, and parts of Oman). Non-literary languages and dialects: Tātī, Tālišī and Gīlakī (on the shores of the Caspian Sea), Central dialects (spoken in a vast area between Hamadān, Kāšān and Iṣfahān), Kirmānī (south of the Dašt-i Kawīr).

1.1.2 Eastern Iranian languages

1.1.2.1 Southeastern group

Pashto (Afghanistan, Pakistan, eastern border region of Iran), Pamir languages (Pamir Mountains along the Pānj River).

1.1.2.2 Northeastern group

Yaɣnōbi (Zarafšān region of Tajikistan), Ossetic (central Caucasus).

1.2 Historical development of Arabic–Persian language contact

Language contact between Arabic and Persian has been a reciprocal process for the past 1500 years. During the pre-Islamic and early Islamic era (sixth–seventh centuries CE), Middle Persian, being embedded in the well-established and sophisticated Iranian culture, provided many loanwords to pre-Classical and Classical Arabic (Gazsi 2011: 1015; see also van Putten, this volume) under RL (recipient-language) agentivity (Van Coetsem 1988; 2000). With the collapse of the Sasanian Empire and expansion of Islam and the Arabic language over vast territories outside Arabia, Classical Arabic began to exercise an unprecedented impact on the emerging New Persian language. Arabic never took root in the everyday communication of the ethnically Persian population, although it gained some dominance as a written vehicle in the administrative, theological, literary and scientific domains in the eastern periphery of the Abbasid Caliphate. Instead, spoken Middle Persian (*Darī*) flourished as a vernacular language. In the middle of the ninth century CE, it was in this part of Iran, specifically in Fārs province, that *Darī* emerged in a new form as it repositioned itself in the culture and literature of the local populace. This new literary language, the revitalization of the Persian linguistic heritage, would be called New Persian. Since its earliest phase, New Persian has borrowed a staggering number of loanwords. Initially, these loanwords were borrowed from various northwestern and eastern Iranian languages, such as Parthian and Sogdian. Despite this relatively large group of loans, the most versatile lenders were the Arabs. Whereas in the pre-Islamic era Arabic had almost exclusively taken lexical items from Middle Persian (in the fields of religion, botany, science and bureaucracy among others), New Persian also incorporated Arabic morphosyntactic elements.

The first Arabic loanwords began to permeate New Persian in the ninth–tenth centuries CE (20–30%). This process was not even diminished by the Iranian *šuʕūbiyya* movement, the major output of which was all conducted in Arabic. In subsequent centuries, Persian continued to absorb an ever-expanding set of Arabic lexemes. By the turn of the twelfth century, the proportion of Arabic loans increased to approximately 50%. The majority of Arabic loans had already been integrated into New Persian by that time and have shown a remarkable steadiness until recently.

After the fall of Baghdad in 1258 CE, Arabic lost its foothold in the eastern provinces of the Caliphate, thereby drawing the final boundary between the use of Arabic and Persian (Danner 2011). The Mongol Ilkhānids, who as non-Muslims were not dependent on Arabic, introduced Persian as the language of education and administration in Iran and Anatolia. Despite the significant de-

struction the Mongols caused to northern Iran during their conquest, this period (thirteenth and foureenth centuries CE) is considered to be the zenith of Persian literature. This is also the epoch when literary Persian is, in an excessive way, inundated with Arabic language elements. This phenomenon is easily detectable in the works of one of the most significant personalities in Classical Persian literature, and a pre-eminent poet of thirteenth-century Persia, Saʕdī of Shiraz. Following the norms of Persian prose writing and poetry of his time, Saʕdī flooded his writings with a bewildering array of Arabic language elements. To illustrate this, here is a typical sentence from Saʕdī's *Gulistān* 'Rose Garden' (completed in 1258 CE), where words of Arabic origin are highlighted in boldface (Yūsifī 2004: 77).

(1) اگر رای عزیز فلان ، أحسن الله خلاصه ، به جانب ما التفات کند در رعایت خاطرش هرچه تمامترسعی کرده شود واعیان این مملکت به دیدار او مفتقرند و جواب این حروف را منتظر
agar ***rāy****-i* ***ʕazīz****-i* ***fulān****,* ***aḥsana allāhu ḫalāṣahu****, ba* ***ǧānib****-i mā* ***iltifāt*** *kunad dar* ***riʕāyat****-i* ***ḫāṭir****aš har či* ***tamām****tar* ***saʕī*** *karda šawad wa* ***aʕyān****-i īn* ***mamlakat*** *ba dīdār-i ū* ***muftaqir****and wa* ***ǧawāb****-i īn* ***ḫurūf*** *rā* ***muntaẓir****.*

'If the precious mind of that person, may God make the end of his affairs prosperous, were to look in our direction, the utmost efforts would be made to please him, because the nobles of this realm would consider it an honor to see him, and are waiting for a reply to this letter.'[1]

It is easy to ascertain that, apart from verbs and adverbs, almost every lexical item in the sentence is of Arabic origin. But writers of this era, such as Saʕdī, not only inundated their works with Arabic elements, but even used Arabic morphology and semantics freely by coining new and innovative meanings, e.g. *ṣaʕqa* 'lightning' < MSA/MSP *ṣāʕiqa* or *baṭṭāl* 'liar' < MSA/MSP 'inactive, unemployed person',[2] < MSA *mubṭil* 'liar'. The Persian and Arabic language use of Saʕdī and other literary figures in the Classical Persian period came closest to what Lucas (2015) calls convergence under the language dominance principle. As reflected in the purely Arabic and Arabic-infused Persian segments of his oeuvre, Saʕdī was equally dominant in both Classical Arabic and Classical Persian along with the dialect of Shiraz.

[1]Persian transcription in this chapter follows the Arabic phonological conventions to avoid using two disparate systems.

[2]In the chapter, references are made to Modern Standard Arabic (MSA) and Modern Standard Persian (MSP) as a comparison to dialectal forms in both languages. This seemed more straightforward as it is not always feasible to ascertain at what point in time a lexeme was borrowed from Arabic into Persian.

Modern Persian is still deeply rooted in Arabic. Arabic loanwords constitute more than 50% of its vocabulary, but in elevated styles (religious, scientific, literary) Arabic loans may exceed 80% (Jeremiás 2011). Although the proportion of these loanwords fluctuates according to age, genre, social context or idiolect, any style in Modern Persian deprived of Arabic influence is almost impossible. An endeavor similar to Atatürk 's to purge Turkish of foreign language elements would be unrealistic in Modern Persian, even with recurring efforts by linguistic purists and the Academy of Persian Language and Literature (*Farhangistān-i zabān wa adab-i fārsī*).[3] It is noteworthy that when the need arose for new terminology to describe fledgling political concepts in Iran, for instance during the Constitutional Revolution in the early twentieth century, as Elwell-Sutton (2011) phrased it, "politicians and journalists instinctively turned to Arabic rather than Persian". Frequently, however, these "Arabic" words were new coinages in the recipient language, e.g. *mašrūṭa* 'constitution', *mawqiʕiyyat* 'situation, position'. After the Islamic Revolution in 1979, another wave of Arabic lexemes related to the new religious governing system surfaced, e.g. *mustaẓʕifīn* 'the needy, the enfeebled' (< MSA *mustaḍʕafūna*/*mustaḍʕafīna*).

Primary and secondary schools in contemporary Arabic-speaking countries do not offer language education in Persian. In Iran, compulsory Classical Arabic instruction is part of the curriculum. However, the language is taught for religious purposes only, with no intention to utilize MSA as a means of acquiring communication skills.

2 Contact languages

This section briefly describes the linguistic impact of Standard Arabic on several New Iranian languages. A more detailed analysis of contact-induced language change in New Persian (*Fārsī*) will follow in §3.

2.1 Arabic influence on New Iranian languages

2.1.1 Tajik (*Tōǧīkī*)

Tajik, written in a modified Cyrillic script, is the variety of New Persian spoken throughout Central Asia, most notably in Tajikistan, Uzbekistan, and northern Afghanistan. Similarly to all varieties of Persian, Arabic borrowings constitute

[3] An example of their activity is the publication (by Rāzī 2004) of a dictionary that lists "pure" Persian words.

the earliest layer of foreign vocabulary in Tajik (Perry 2009). This lexicon was transferred under RL agentivity. Although Arabic lexical items have a firm hold in Tajik, their pattern of distribution differs from that of New Persian. For instance, Tajik uses *pēš* 'before' and *pas* 'after' rather than MSA/MSP *qabl* and *baʕd*, but *ōid ba-/ōid-i* (< MSA *ʕāʔid* 'returning') 'concerning, relating to' in lieu of MSP *rāǧiʕ ba-* (< MSA *rāǧiʕ* 'recurring'). Also, *madaniyyat* 'civilization' (< MSA *madaniyya* 'civilization'; cf. MSA/MSP *tamaddun* 'civilization'), *hōzir* 'now' (< MSA *ḥāḍir* 'present; ready', MSP *ḥāẓir* 'present'), *ittifōq* '(labor) union' (< MSA *ittifāq* 'agreement; contract'; cf. MSP *ittiḥād* '[labor] union').

Arabic plural forms, both sound feminine plural and broken plural, were lexicalized with collective or singular meanings: *hašarōt* 'insect', with regular plural ending *hašarōthō* 'insects' (< MSA/MSP *ḥašarāt* 'insects'), *talaba* 'student', pl. *talabagōn* (< MSA/MSP *ṭalaba* 'students'), *šarōit* 'condition, stipulation' (< MSA/MSP *šarāʔiṭ* 'conditions').

2.1.2 Kurdish

A characteristic feature of Kurdish, the change of postvocalic /m/ > /v/ or /w/, also occurs frequently in words of Arabic origin: *silāv* 'greeting' (< MSA/MSP *salām*; Paul 2008).

2.1.3 Gurānī

The phonological system of Gurānī dialects is similar to Kurdish in the occurrence of Arabic pharyngeal and emphatic sounds /ʕ/, /ḥ/, /ṣ/ (MacKenzie 2012).

2.1.4 Ossetic

Ossetic has incorporated terms related to Islam from Arabic and Persian through neighboring Caucasian languages (Thordarson 2009).

2.2 Arabic-speaking communities in Iran

Arabic-speaking communities are known to be present within the boundaries of the Islamic Republic of Iran, but their exact number is not readily discernible from official statistics. It is estimated that 3% of Iran's 80 million citizens are Arabs, which would put the Arab population at approximately 2.5 million. The majority of Arabs live in the western parts of Khuzestan province (see Leitner, this volume), but also along Iran's Persian Gulf coast and parts of Khorasan in eastern Iran (Oberling 2011). Already during the Sasanian era, several Arab tribes,

including the Bakr ibn Wāʔil and Banū Tamīm, settled in the area stretching from the Šaṭṭ al-ʕArab to the Zagros Mountains (Daniel 2011). At the end of the sixteenth century, the Banū Kaʕb, originating from present-day Kuwait, settled in Khuzestan. During subsequent centuries, more Arab tribes moved from southern Iraq to the province. As a result, Khuzestan, which until 1925 was called ʕArabistān, became extensively Arabized. Members of these Arab tribes live on either side of the Iran–Iraq border. In the same way as Iraqi Arabic vernaculars, Khuzestan Arabic has been influenced by Persian. However, Khuzestan Arabic can most easily be distinguished from Iraqi dialects by its wide-ranging transfer of Persian lexemes (Ingham 1997: 25; see also Leitner, this volume).

Arab presence has a well-documented history on the Iranian coastline of the Persian Gulf, in what now constitutes Būšihr and Hurmuzgān province. According to travelogues from the eighteenth to twentieth centuries CE, as well as British archival materials dating back to the British Residency of the Persian Gulf, Arab tribes inhabited most fishing and pearling villages, as well as islands and coastal towns with strategic importance (e.g. Bandar ʕAbbās). The most significant tribes in this area were, and in some cases still are, the Qawāsim, Marāzīq, Āl Ḥaram, Āl ʕAlī, Āl Naṣūr, Banī Tamīm, Banī Ḥammād, Banī Bišr, among others. In contrast to most Persians and Khuzestani Arabs who are primarily Shiite, these tribes are Sunni Muslims. A widespread exonym to designate Arabs on the Iranian coast, but shunned by the local population, is *hōla* (variously referred to as *hula*, *huwala* or *hawala*). Local tribes prefer the endonym 'Arabs of the Coast' (*ʕarab as-sāḥil*) (Gazsi 2017: 110).

Most Khuzestani and Iranian Persian Gulf Arabs are bilingual, speaking Arabic as their mother tongue and Persian as a second language. Although Khuzestan and the two Persian Gulf provinces are geographically part of Iran, linguistically their Arab populations form a continuum with the southern Mesopotamian Muslim *gilit* dialects, and the dialects of the eastern coast of the Arabian Peninsula, respectively. In the spoken and written code, 'Arabs of the Coast' often engage in tetra-glossic switching between MSA, Gulf Arabic (GA), MSP, Colloquial Persian and one of its local dialects such as Bandarī. In their speech, Persian phonological and lexical elements are borrowed into GA under RL agentivity.

3 Contact-induced changes in New Persian and modern Persian dialects

Language contact between Arabic and New Persian is most evidently detectable in the New Persian lexicon, and to a lesser extent in phonology and morphosyn-

tax. This section summarizes the characteristics of this contact. In addition to standard New Persian, and its contemporary variant MSP, Arabic has also influenced modern Persian dialects. This influence is slightly different, and in several ways more far-reaching, particularly in the realm of phonology and lexicon.

Persian dialects developed separately from and parallel to Classical Persian and MSP. Modern Persian dialects retain several Early Classical and Classical Persian phonological and morphosyntactic features that are not present in MSP. Additionally, they were in direct contact with the Arabic language through Arab tribes that settled across Persia immediately after the Islamic conquest or in later centuries. Although most Arab tribes have long been integrated into the Persian-speaking population, the Arabic language in the areas currently dominated by ethnic Arabs is still in contact with the surrounding Persian dialects. Unlike Arabic influence on the standard version of New Persian, Arabic influence on modern Persian dialects is an understudied field that does not allow for providing an exhaustive list of contact-induced changes at this point. Instead, below is a preliminary description of salient examples of how Arabic phonological and lexical elements were transferred to New Persian, both its standard and dialectal variations.

3.1 Phonology

3.1.1 New Persian

The initial step in the adoption of Arabic lexemes was the adoption of the Arabic script. New Persian began to use a modified Arabic script in the ninth century CE; it has 32 letters, 28 taken over from Arabic and 4 new letters added to represent Persian phonemes (/p/, /č/, /ž/, /g/). Arabic /θ/ and /ṣ/ collapse to Persian /s/, Arabic /ð/, /ḍ/, /ẓ/ collapse to Persian /z/, and Arabic /ṭ/ becomes Persian /t/. The phonemic inventory of Early Classical Persian was augmented with the glottal stop, which originated in the two separate Arabic phonemes /ʔ/ and /ʕ/.

3.1.2 Modern Persian dialects

This section highlights phonological features of modern Persian dialects that were the result of contact-induced language change under RL agentivity, either with Arabic or with Classical Persian and subsequently MSP.

3.1.2.1 Adoption of Arabic pharyngeal sounds

The two Arabic pharyngeal sounds undergo phonological integration in New Persian: the voiceless pharyngeal fricative /ḥ/ is pronounced as a voiceless glottal fricative /h/, and the voiced pharyngeal fricative /ʕ/ as a glottal stop /ʔ/. The dialects of Dizfūl and Šūštar have acquired pharyngeal sounds directly from Arabic, which occur in Arabic loanwords: *ʕağīb* 'strange', *baʕd* 'after' (MacKinnon 2015). The dialect of Jarkūya shares this feature: *ḥasüd* 'jealous', *ğimʕa* 'Friday' (Borjian 2008).

The dialect of Kulāb in Tajikistan also borrows Arabic pharyngeal sounds in words of Arabic origin: *ʕaib* 'flaw', *daʕvō* 'claim', *mıʕalim* 'teacher', *ḥıkımat* 'wisdom', *sōḥib* 'owner'. Arabic pharyngeal sounds also occur in a few Persian/Tajiki words (*ʕasp* 'horse', *ḥamsōya* 'neighbor'). Interestingly, the pharyngealized form for 'horse' occurs far and wide within the Iranian linguistic domain, as *ʕasb* in the Lurī dialect of Šūštar, in Ḫānsāri and Caucasian Tātī. In the Arab Gulf states, the *ʕAğam*, ethnic Persians holding Kuwaiti, Emirati and other Gulf citizenship, pronounce Arabic loanwords in their Persian speech with pharyngeal sounds.

3.1.2.2 Dropping of Arabic pharyngeal sounds

In several modern Persian dialects, the voiceless pharyngeal fricative /ḥ/ is absent. The preceding vowel is lengthened or the subsequent vowel disappears too, e.g. *mūtāğ* 'in need, destitute' < MSA/MSP *muḥtāğ* (Īzadpanāh 2001: 190), *ṣārā* 'desert' < MSA/MSP *ṣaḥrā* (Sarlak 2002: 15), *ṣāb* 'owner' < MSA/MSP *ṣāḥib* (Ṣarrāfī 1996: 135), *mulāẓa* 'consideration, observation' < MSP *mulāḥiẓa*, cf. MSA *mulāḥaḍa* (Ṣarrāfī 1996: 188), *ṣul* 'peace' < MSA/MSP *ṣulḥ* (Stilo 2012), *ēsās* 'feeling' < MSA/MSP *iḥsās* (Salāmī 2004: 160–161). In Kirmān, the sound change /uḥ/ > /ā/ is attested, e.g. *fāš* 'insult' < MSA/MSP *fuḥš* (Borjian 2017).

The voiced pharyngeal fricative /ʕ/, pronounced as a glottal stop in MSP, can also be dropped. This may result in vowel lengthening: *māṭal* 'idle' < MSA/MSP *muʕaṭṭal* (Ṣarrāfī 1996: 184), *māmila* 'transaction' < NewP *muʕāmila*, cf. MSA *muʕāmala* (Ṣarrāfī 1996: 184; Sarlak 2002: 15), *rubbi sāt* 'quarter hour' < MSP *rubʕ sāʕat*, cf. MSA *rubʕ sāʕa* (Ṣarrāfī 1996: 108), *mānī* 'meaning' < MSP *maʕnī*, cf. MSA *maʕnā* (Sarlak 2002: 15), *mōğiza* 'miracle' < MSA/MSP *muʕğiza* (Īzadpanāh 2001: 190), *tāğub* ~ *tāğuv* 'surprise, wonder' < MSA/MSP *taʕağğub* (Salāmī 2004: 162–163), *rāyat* 'regard' < MSP *riʕāyat*, cf. MSA *riʕāya* (Ṣarrāfī 1996: 107).

3.1.2.3 Dropping of the Arabic voiceless glottal fricative /h/

The voiceless glottal fricative disappears in closed syllables in many Persian dialects, resulting in occasional vowel lengthening: *ṭārat* 'cleanliness' < MSP *ṭahārat*, cf. MSA *ṭahāra* (Sarlak 2002: 76), *nāal* 'impolite' < MSP *nāahl* (Īzadpanāh 2001: 192).

3.1.2.4 Miscellaneous sound changes

A range of additional consonant developments and shifts can be attested in Persian dialects. Some of these developments include:

/ʕ/ > /ḥ/: In Lurī and the dialect of Jarkūya, a shift occurs from the voiced to the voiceless pharyngeal: *ḥilāğ* 'cure' < MSA/MSP *ʕilāğ* (Īzadpanāh 2001: 207), *ṭaḥna* 'sarcasm' < MSA/MSP *ṭaʃna* (Borjian 2008).

/ḥ/ > /ʔ/ occurring with occasional metathesis: *ṭaʔr* 'plan' < MSA/MSP *ṭarḥ* (Ṣarrāfī 1996: 137), *maʔla* 'city quarter' < MSA/MSP *maḥalla* (Ṣarrāfī 1996: 188), *maʔala* 'city quarter' (Nağībī Fīnī 2002: 133).

/h/ > /ʔ/: *muʔlat* 'deadline, respite' < MSP *muhlat*, cf. MSA *muhla* (Ṣarrāfī 1996: 190).

/θ/ > /t/: This shift is also common in several Arabic dialects, e.g. in Egypt and Morocco: *mīrāt* 'heritage' < MSP *mīrās*, cf. MSA *mīrāθ* (Īzadpanāh 2001: 190).

Word-final /b/ and /f/ > /m/: *nağīm* 'noble' < MSA/MSP *nağīb* (Īzadpanāh 2001: 193), *niṣm* 'half' < MSA/MSP *niṣf* (Īzadpanāh 2001: 195).

/r/ > /l/: in Kirmān, *zilar* ~ *zilal* 'damage, loss' < MSP *ẓarar*, cf. MSA *ḍarar* (Ṣarrāfī 1996: 136; Dānišgar 1995: 163), *ḥaṣīl* 'straw mat' < MSA/MSP *ḥaṣīr* (Ṣarrāfī 1996: 85), *qulfa* 'small room for summer resting' < MSA *ɣurfa* 'room' (Fāẓilī 2004: 151).

Arabic voiceless dental emphatic /ṭ/ > /d/: *mudbaḫ* ~ *madbaḫ* 'kitchen' < MSA *maṭbaḫ* (Ṣarrāfī 1996: 186; not attested in MSP), *mudbaq* in Baḫtiārī (Sarlak 2002: 251).

/b/ > /f/: *muftilā* 'afflicted' < MSP *mubtilā*, cf. MSA *mubtalā* (Borjian 2017).

Medial and word-final /b/ > /v/: in Baḫtiārī, *ādāv* 'customs' < MSA/MSP *ādāb* (Sarlak 2002: 15), *ʕajīv* 'strange' < MSA/MSP *ʕajīb* (Sarlak 2002: 25), *qavīla* 'tribe' < MSA/MSP *qabīla* (Sarlak 2002: 199).

Word-initial /ḫ/ > /h/: in northern Lurī and Baḫtiārī, *hāla* 'aunt' < MSA/MSP *ḫāla* (Īzadpanāh 2001: 204).

/q/ > /k/: *kabīla* 'tribe' < MSA/MSP *qabīla* (Nağībī Fīnī 2002: 21).

/ɣ/ > /q/: *šuql* 'occupation' < MSP/MSA *šuɣl* (Stilo 2012).

/ğ/ > /y/: direct borrowing from Khuzestan Arabic dialects, *mailis* 'council' < MSA/MSP *mağlis* (Sarlak 2002: 260; Fāẓilī 2004: 165).

Metathesis: *qulf* 'lock' < MSA/MSP *qufl* (Salāmī 2004: 84–85; Imām Ahwāzī 2000: 146), *ṣuḥb* 'morning' < MSA/MSP *ṣubḥ* (Dānišgar 1995: 161; Nağībī Fīnī 2002: 23).

The full /t/ of the *tāʔ marbūṭa* appears on words where it is absent in MSP: *ḥalmat* 'attack' < MSA/MSP *ḥamla* (Īzadpanāh 2001: 207), *ḥağāmat* 'cupping' < MSA/MSP *ḥağāma* (Salāmī 2004: 92–93). This was a typical feature of Classical Persian literature.

3.2 Morphosyntax

Several Arabic morphosyntactic features were transferred to New Persian in the realm of nominal morphology under RL agentivity. These features encompass sound and broken plural forms (*musāfirīn* 'passengers', *tablīɣāt* 'propaganda', *dihāt* 'villages', *ḥuqūq* 'rights'), possessive constructions (*fāriɣ ut-taḥṣīl* 'graduate', *wāğib ul-iğrā* 'peremptory') and occasional gender agreement in lexicalized expressions (*quwwa-yi darrāka* 'perceptive power'). Word formation has been an active method of transferring Arabic lexical elements into New Persian from early on, either by way of derivation (*diḫālat* 'interference' < MSA *mudāḫala*, *awlā-tar* 'superior' < MSA *awlā*, *raqṣīdan* 'to dance', *akṯaran* 'most, generally' < MSA *akθar* 'more, most') or compounding. Compounding is a highly developed process of enlarging the New Persian vocabulary. It is manifest in lexical compounds (*tayẓia-šināṣ* 'nutritionist', *ḫiānat-kārāna* 'perfidiously') and phrasal compounds (*iṭāʕat kardan* 'to obey', *ʕadam-i wuğūd* 'non-existence', *ʕala l-ḫuṣūṣ* 'particularly').

3.3 Lexicon

3.3.1 Arabic lexicon in New Persian

Contact-induced language change manifests itself most strikingly in the lexicon transferred from Arabic to New Persian under RL agentivity. The earliest loanwords entered New Persian during the ninth–tenth centuries. This process occurred smoothly, as the phonological inventory of Early Classical Persian was likely close to that of Middle Persian and also close to that of Classical Arabic.[4] The influx of Arabic loanwords has unabatedly continued over the centuries until now. To showcase a recent example of Arabic vocabulary in Modern Persian, below are titles of articles from *Hamšahrī* 'fellow citizen', a major Iranian national newspaper, taken from its 29th January 2018 edition. Arabic words are highlighted in boldface:

(2) a. **kulliyyāt**-i **lāyiḥa**-yi būdğa-yi sāl-i 97-i **kull**-i
total.PL-GEN bill-GEN budget-GEN year-GEN 97-GEN whole-GEN
kišwar **radd** šud
country reject be.PST.3SG
'The total budget bill of the year 1997 for the whole country was rejected.'

b. **daʕwat** az tihrānīhā barā-yi **ihdā**-yi ḫūn **asāmī**-yi
call from Tehrani.PL for-GEN donation-GEN blood name.PL-GEN
marākiz-i **faʕʕāl**
center.PL-GEN active
'Calling the residents of Tehran to donate blood. Names of active centers.'

c. **iḥrāz**-i **huwiyyat** dar **muʕāmilāt**-i **milkī**
authentication-GEN identity in transaction.PL-GEN proprietary
bā kārt-i hūšmand-i **millī** anğām mī-šaw-ad
with card-GEN smart-GEN national complete PRS-be-3SG
'Personal authentication in real estate transactions is done with the national smart card.'

In the Arabic lexicon of New Persian, further characteristics can be observed, such as phonetic changes (NewP *maʔnī* 'meaning' < MSA *maʕnā*, NewP *madrisa* 'school' < MSA *madrasa*, NewP *šikl* 'shape, form' < SA *šakl*), where in some cases

[4] In Early Classical Persian, short vowels were likely pronounced as /u/ and /i/, and the *alif* as /ā/. In MSP, the pronunciation is /o/, /e/ and /ɒ/.

the Persian pronunciation may follow Arabic dialectal forms, semantic changes (NewP *kitābat* 'writing' and *kitāba* 'inscription' < MSA *kitāba* 'writing', NewP *ṣuḥbat* 'speech' < MSA *ṣuḥba* 'companionship'), and occasional *imāla* in elevated or poetic style (NewP *ḥiğīz* < MSA *ḥiğāz*).

3.3.2 Arabic lexicon in Persian dialects

Arabic loanwords affect Persian dialects in two ways that differ from MSP: i) semantic changes, where Arabic lexemes assume new meanings unattested in both MSA and MSP: in Kirmān *ðāt* 'age' (Ṣarrāfī 1996: 106) < MSA/MSP 'self, soul, essence, nature', *ðātī* 'old' < MSA/MSP 'own, personal'; ii) lexemes and expressions directly borrowed from Arabic, and not attested in MSP: in Šūštar, *ḥaya* 'snake' < MSA *ḥayya*, MSP *mār* (Fāẓilī 2004: 140), *ṭayyāra* 'airplane' < Arabic dialects *ṭayyāra*, MSA *ṭāʔira*, MSP *hawāpaimā* (Fāẓilī 2004: 150), *ṣaḥn* 'bowl, dish' < MSA *ṣaḥn*, MSP *bušqāb* (Fāẓilī 2004: 150), *ṭabaq* 'plate, tray' < MSA *ṭabaq*, MSP *sīnī* (Fāẓilī 2004: 150), in Fīn, *mismāl* 'nail' < MSA *mismār*, MSP *mīḫ* 'nail' (Nağībī Fīnī 2002: 133), in Kirmān, *aḥad un-nās* 'nobody, somebody' < MSA *aḥad un-nās*, MSP *hīčkas* 'nobody', *kasī* 'somebody' (Ṣarrāfī 1996: 33).

On the Persian Gulf coast of Iran, due to linguistic, economic and commercial connections with the Arabian Peninsula, Persian dialects have incorporated a number of Arabic technical terms relating to pearling, fishing and traditional shipbuilding from Gulf Arabic: *muḥār* 'shellfish, oysters' (cf. MSA *maḥār*), *giyās* 'measure, gauge' (< GA *giyās*, cf. MSA *qiyās*), *mīdāf* 'helm (boat)' (< GA *mīdāf*, cf. MSA *miğdāf*), *māčila* 'meal (on a boat)' (< GA *māčila*, cf. MSA *maʔkūl*). Two neighborhoods in the town of Bandar Linga (opposite Dubai, 180 km west of Bandar ʕAbbās) are called *Maḥalla-yi Baḥrainī* 'Bahraini Quarter' and *Maḥalla-yi Sammāčī* 'Fishers' Quarter' (< GA *sammāč*, cf. MSA *sammāk*) (Baḫtiyārī 1990: 137–138).

4 Conclusion

Although Arabic–Persian language contact has been a well-known phenomenon for centuries, academic research dedicated to this topic is far from abundant. Throughout the centuries, Persian writers and poets used Arabic lexical elements in new meanings or coined non-standard Perso-Arabic lexemes based on Arabic derivational patterns. Idiosyncratic features of individual Persian writers should be examined separately before compiling a comprehensive review of this contact-induced language change. Substantial fieldwork needs to be conducted to de-

scribe the bilingualism of ethnic Arab communities of Iran and ethnic Persians in Arabic-speaking countries. Additionally, it is essential for linguists to look into Arabic influence on Modern Persian dialects and Iranian languages other than New Persian. This will help scholars understand the scale and depth of how Arabic has shaped Iranian languages for the past thousand years.

Contact-induced language change in New Iranian languages primarily transpired under RL agentivity. It should be noted, however, that medieval Persian literati were so well-versed in Arabic due to its prestige and dominance, that their bilingualism may have enabled convergence in Arabic–Persian language contact.

Further reading

Asbaghi (1987) gathers eight hundred Persian words of Arabic origin in twenty-three groups and analyzes the semantic changes they underwent when transferred from Arabic to New Persian.
Gazsi (2011) gives an overview of Arabic–Persian language contact from pre-Islamic times up to the modern era, also touching on Arabic dialects in Iran. It additionally provides a brief analysis of Arabic morphosyntactic features in New Persian.
Ṣādiqī (2011) discusses a range of Arabic phonological, grammatical and semantic elements in New Persian.

Acknowledgements

I would like to express my gratitude to Prof. Éva Jeremiás, Prof. Werner Arnold and Prof. Ali Ashraf Sadeghi for their support while I was working on Arabic–Persian language contact. I am thankful to members of the 'Arabs of the Coast' in the UAE, especially Sheikh Ibrahim, Sheikh Abdulrahman, Dr. Abdullah, Walid, Ahmed, and many others for providing language data in GA.

Abbreviations

BCE	before Common Era
CE	Common Era
GA	Gulf Arabic
GEN	genitive
MSA	Modern Standard Arabic
MSP	Modern Standard Persian
NewP	New Persian
PL	plural
PRS	past
PST	past
RL	recipient language

References

Asbaghi, Asya. 1987. *Die semantische Entwicklung arabischer Wörter im Persischen*. Stuttgart: Steiner.

Baḫtiyārī, Majīd. 1990. *Ustān-i hurmuzgān (rāhnamā-yi mufaṣṣal-i Īrān)*. Tehran: Gītāšināsī.

Borjian, Habib. 2008. Jarquya ii: The dialect. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Borjian, Habib. 2017. Kerman xvi: Languages. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Daniel, E. L. 2011. ʿArab iii: Arab settlements in Iran. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Dānišgar, Aḥmad. 1995. *Farhang-i wāžahā-i rāyiğ-i turbat-i ḥaidarīya*. Mašhad: Āstān-i Quds-i Raẓawī.

Danner, Victor. 2011. Arabic language iv: Arabic literature in Iran. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Elwell-Sutton, L. P. 2011. Arabic language iii: Arabic influences in Persian literature. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Fāẓilī, Muḥammad Taqī. 2004. *Farhang-i gūyiš-i šūštarī*. Tehran: Pāzīnah.

Gazsi, Dénes. 2011. Arabic–Persian language contact. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1015–1021. Berlin: De Gruyter Mouton.

Gazsi, Dénes. 2017. Language and identity among the 'Arabs of the coast' in Iran and the Arab Gulf states: Social media and beyond. *RiCognizioni* 7. 105–129.

Imām Ahwāzī, Sayyid Muḥammad ʿAlī. 2000. *Čīstānnāma-yi dizfūlī*. Tehran: Daftar-i Pažūhišhā-yi Farhangī.

Ingham, Bruce. 1997. *Arabian diversions: Studies on the dialects of Arabia*. Reading: Ithaca Press.

Īzadpanāh, Ḥamīd. 2001. *Farhang-i lurī*. Tehran: Intišārāt-i Asāṭīr.

Jeremiás, Éva M. 2004. New Persian. In Peri Bearman, Thierry Bianquis, C. Edmund Bosworth, Emeri Johannes van Donzel & Wolfhart P. Heinrichs (eds.), *Encyclopaedia of Islam*, 2nd edn., vol. 12: *Supplement*, 426–448. Leiden: Brill.

Jeremiás, Éva M. 2011. Iran. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Lucas, Christopher. 2015. Contact-induced language change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.

MacKenzie, David N. 2012. Gurāni. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

MacKinnon, Colin. 2015. Dezfūlī and Šūštarī dialects. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Nağībī Fīnī, Bihğat. 2002. *Barrasī-yi gūyiš-i fīnī*. Tehran: Našr-i Āṣār-i Farhangistān-i Zabān wa Adab-i Fārsī.

Oberling, Pierre. 2011. ʿArab iv: Arab tribes of Iran. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Paul, Ludwig. 2008. Kurdish language i: History of the Kurdish language. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Perry, John R. 2009. Tajik ii: Tajik Persian. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Rāzī, Farīda. 2004. *Farhang-i wāžahā-yi fārsī-yi sara barā-yi wāžahā-yi ʃarabī dar fārsī-yi muʃāṣir*. Tehran: Našr-i Markaz.

Ṣādiqī, ʕAlī Ašraf. 2011. Arabic language i: Arabic elements in Persian. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Salāmī, ʕAbdulnabī. 2004. *Ganǧīna-yi gūyiššināsī-yi fārs*. Tehran: Našr-i Āṣār-i Farhangistān-i zabān wa adab-i Fārsī.

Sarlak, Riẓā. 2002. *Wāžanāma-yi gūyiš-i baḫtiārī-yi čahārlang*. Tehran: Āṣār.

Ṣarrāfī, Maḥmūd. 1996. *Farhang-i gūyiš-i kirmānī*. Tehran: Surūš.

Schmitt, Rüdiger. 1989. *Compendium linguarum iranicum*. Wiesbaden: Reichert.

Skjærvø, Prods Oktor. 2012. Iran vi: Iranian languages and scripts. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Stilo, Donald. 2012. Gīlān x: Languages. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Thordarson, Fridrik. 2009. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

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.

Yūsifī, Ɣulāmḥusain. 2004. *Gulistān-i saʕdī*. Tehran: Širkat-i Sahāmī-yi Intišārāt-i Ḫārazmī.

Chapter 8

Kurdish

Ergin Öpengin
University of Kurdistan-Hewlêr

This chapter provides an overview of the influence of Arabic on Kurdish, especially on its Northern and Central varieties spoken mainly in Turkey–Syria–Iraq and Iraq–Iran, respectively. It summarizes and critically assesses the limited research on the contact-induced changes in the phonology and syntax of Kurdish, and proposes several new dimensions in the morphology and syntax in addition to providing also a first treatment of lexical convergence in Kurdish through borrowings from Arabic.

1 Kurdish and its speech community

Kurdish is a Northwestern Iranian language spoken by 25 to 30 million speakers in a contiguous area of western Iran, northern Iraq, eastern Turkey and northeastern Syria. There are also scattered enclaves of Kurdish speakers in central Anatolia, the Caucasus, northeastern Iran (Khorasan province) and Central Asia, with a large European diaspora population. The three major varieties of Kurdish are: (i) Southern Kurdish, spoken under various names near the city of Kermanshah in Iran and across the border in Iraq; (ii) Central Kurdish (also known as Sorani), one of the official languages of the autonomous Kurdish region in Iraq, also spoken by a large population in western Iran along the Iraqi border; (iii) Northern Kurdish (also known as Kurmanji), spoken by the Kurds of Turkey, Syria and the northwestern perimeter of Iraq, in the province of West Azerbaijan in northwestern Iran and in pockets in the west of Armenia (cf. Haig & Öpengin 2014 for a discussion on defining "Kurdish"). Of these three, the largest group in terms of speaker numbers is Northern Kurdish. The Kurdish population in respective states is difficult to reliably determine since none of the sovereign countries make the relevant census information available. Table 1 provides some

Ergin Öpengin. 2020. Kurdish. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 175–201. Berlin: Language Science Press. DOI:??

cautious estimates based on various sources (especially Sirkeci 2005; Zeyneloğlu et al. 2016; and Ethnologue).[1,2]

Table 1: Estimates of Kurdish population numbers

Country	Population size
Turkey	c. 15,000,000
Iraq	c. 6,000,000
Iran	c. 8,000,000
Syria	c. 2,000,000

2 The history of Kurdish–Arabic contact

Information about the pre-Islamic history of the Kurds and their language is scarce. According to early Islamic sources, at the time of the Islamic conquest of the Near East (Upper Mesopotamia, Iran, and Armenia) in the seventh century (Bois et al. 2012: 451), the communities designated with the term *Kurd* were already living in most of the present-day Kurdish-inhabited areas, namely from Mosul to the north of Lake Van, and from Hamadan to the Jazira region situated around the intersection of present-day Syria, Iraq and Turkey (James 2007: 111). The Kurds have thus been living in contact with various Aramaic-speaking Christian and Jewish communities as well as Arabic-speaking communities since at least the early Islamic period, though the contact of Iranian-speaking populations with Aramaic dates back to the fifth century BCE (cf. Utas 2005: 69, citing also Folmer 1995 and Kent 1953). Kurdish differs from other Iranian languages such as Persian in sharing the same or close geographical spaces with Arabic-speaking populations, especially in Upper Mesopotamia. The historical socio-cultural contact between Kurdish and Arabic-speaking communities requires a more refined treatment than is currently possible, but there are a number of medieval Arabic sources which attest to the interaction and mobility of Kurdish and Arabic communities in some regions (e.g. Erbil, Mosul), as well as language shift of some Kurdish communities to Arabic and vice versa (cf. Bois et al. 2012: 449, 452, 456; James 2007: 115–120).

[1]See https://www.ethnologue.com/language/kur.

[2]The population figures should not be taken as equivalent to "number of speakers", since especially in Turkey a significant portion of the Kurdish population grow up with no or very limited knowledge of Kurdish (cf. Öpengin 2012; Zeyneloğlu et al. 2016).

Given the unquestionably prestigious status of Arabic in administration and sciences in the Islamicized Near East, consolidated especially under Abbasid rule (which included most of the Kurdish-inhabited areas), Kurdish was heavily dominated by Arabic. Even in several of the important medieval Kurdish dynasties such as that of the Marwānids (10th–11th centuries), Arabic enjoyed the high status of being the administrative and literary language (cf. James 2007: 112), since the coins bore Arabic script, while *qaṣīda* reading ceremonies or contests would feature primarily Arabic, but to a limited extent also Persian pieces (Ripper 2012: 507–528). With the conquest of the Kurdish-inhabited regions by Turkic peoples and Mongols from tenth century onwards, which led also to the final overthrowing of the Abbasid state in 1248 by the Mongols, the Arabic-speaking populations may have started to diminish and retreat. Although at this stage Persian attained a firm status as the literary language in the Islamic East (Perry 2012: 73), Arabic preserved its higher status in administration and, later on, especially in education, well into the end of the nineteenth century. Thus, Kurdish developed a literary tradition only starting from the sixteenth century, but its limited usage was largely restricted to writing verse throughout the following several centuries. The literature in this period is heavily dominated by the vocabulary and literary formulas and metaphors of the two dominant languages, Arabic and Persian (cf. Öpengin forthcoming).

In the early twentieth century, with the dissolution of the Ottoman Empire, Kurdish in Iraq and Syria again came into primary direct contact with Arabic. In Iraq, up until 1991, with the establishment of a Kurdish autonomous region, the language configuration was one in which Arabic was the prestigious language of higher domains. Not being in possession of any official status, the Kurds in Syria have been in a highly asymmetric language-contact situation with Arabic. In Turkey, especially in Mardin and Siirt provinces, Kurds have been in contact with Arabic-speaking communities, but as the lingua franca of the communities of cultural–historical Kurdistan (cf. Edwards 1851: 121), Kurdish must have been the dominant language of interaction between these communities, and it is indeed possible to observe important influences from Kurdish on the local Arabic dialects (cf. Jastrow 2011 and §3.1 below.).

As a result of these differing degrees and modalities of contact with Arabic, the influence of Arabic should be viewed as consisting of at least two layers, and viewed separately for different country contexts where Kurdish is spoken. Of the two layers, there should be assumed a deeper contact influence, shared in larger portions of Kurdish-speaking areas, dating to before the twentieth century; and a more shallow layer that is the result of the more recent societal bilin-

gualism in Iraq and Syria. Likewise, while in Syria and Iraq the Arabic influence on Kurdish continues, this influence is largely replaced by influence from the dominant state languages in Turkey and Iran. Naturally, the intensity of Arabic influence on Kurdish shows a great deal of variation across Kurdish varieties and dialects within varieties. Accordingly, the historically deeper-layer Arabic influence on Kurdish is characterized by its being restricted mostly to lexicon and being shared in the majority of Kurdish dialects. This has been the result of borrowing under RL-agentivity in the sense of Van Coetsem (1988; 2000). On the other hand, the relatively advanced Arabic influence on the Kurdish spoken in the historical Jazira region (including Mosul, northeast Syria, and Mardin province in southeast Turkey), as well as the more recent Arabic influence on the Kurdish spoken in Syria, but also – albeit more restrictedly – in Iraq, concerns also grammatical constructions and at least some of that contact influence could be due to imposition under SL-agentivity.

3 Contact-induced changes in Kurdish

3.1 Phonology

The consonant inventory of Kurmanji is given in Table 2.[3]

In cells of doublets/triplets, the voiceless phonemes come first. The apostrophe on plosive and fricative phonemes indicates aspiration, which marks a phonemic distinction in Kurmanji. In addition to these consonants with indisputable phonemic status, there are the so-called emphatic or pharyngealized variants of the obstruents /p, b, t, d, s, z/. These variants are transcribed in the text with a dot beneath the characters.

The consonant inventory of Sorani is basically identical with Table 2, except: (i) it does not have unaspirated stop phonemes; and (ii) it has velar nasal and velarized lateral phonemes (Öpengin 2016: 27).

Arabic (or more generally Semitic) influence on the phonology of Kurdish is most clearly observed in the presence of the two pharyngeal phonemes *ḥ* [ħ] and *ʿ* [ʕ] (cf. Kahn 1976; Haig 2007; Anonby forthcoming; Barry forthcoming), as well as the series of emphatic obstruants *ṭ, ḍ, ṣ,* and *ẓ* (Haig & Öpengin 2018), respectively. The precise Semitic source language for these sounds cannot be determined, since Kurdish (or rather its ancestor languages) must have been in close

[3]Kurdish data are transcribed in the standard Kurdish Latin alphabet with some additions for emphatics and pharyngeals, mostly consonant with the Library of Congress approach for the romanization of Kurdish: https://www.loc.gov/catdir/cpso/romanization/kurdish.pdf.

Table 2: Consonant phonemes in Kurmanji

	bilabial	labio-dental	alveolar	palatal	velar	uvular	pharyngeal	glottal
Plosive	p’ p b		t’ t d		k’ k g	q		’
Fricative		f v	s z	ş j	x ẍ		ḥ ʿ	h
Affricate				ç’ ç c				
Nasal	m		n					
Trill			r̄					
Flap			r					
Approximant	w			y				
Lateral			l					

contact with various Semitic languages for more than two millennia (Utas 2005: 69). However, these phonemes set the consonant inventory of Kurdish clearly apart from other West Iranian languages such as Persian, with the only other West Iranian languages possessing both pharyngeals and emphatic consonants being Zazaki, and the Kumzari language spoken mainly in Oman (Anonby forthcoming). In what follows, I illustrate the presence and interactions of the pharyngeal and emphatic consonants in Kurdish, and provide a brief discussion of their paths of development.[4]

The pharyngeal phonemes are found in varying degrees in both Central Kurdish and Northern Kurdish. They are retained in most of the Arabic loanwords originally bearing them, a list of which is given in Table 3.[5]

Some loanwords with original pharyngeals are reanalysed as containing their

[4]The Kurmanji lexical items presented in this section are based on my native-speaker knowledge of the Şemdînan (Şemdinli) dialect, and my knowledge of Kurmanji-internal dialectal variation, drawing also on (Chyet 2003), (Öpengin & Haig 2014), and the Manchester Database of Kurdish Dialects presented in (Matras & Koontz-Garboden 2016). The Sorani lexical items are from Öpengin (2016) and the popular press.

[5]Note that all through the article, unless stated otherwise, the Arabic data represents Classical Arabic, giving an approximation of the ultimate Arabic etyma of the items without necessarily implying that these are the immediate source of the Kurdish items (as they may have been borrowed from local Arabic dialects as well as through the intermediary languages such as Persian or Ottoman). Furthermore, the glosses in tables are for Kurdish items, as sometimes the meanings of the Arabic etyma are not completely identical.

Table 3: Loanwords with retained pharyngeals in Kurdish

Arabic	Northern Kurdish	Central Kurdish	Gloss
ʕarab	*ʿereb*	*ʿereb*	'Arab'
maʕlūm	*meʿlûm*	*meʿlûm*	'evident'
ʕadāla(t)	*ʿedalet*	*ʿedaḷet*	'justice'
ṭābiʕ	*tabiʿ/ṭabiʿ*	*ṭabiʿ*	'dependent'
maḥall	*miḥele*	*meḥel*	'neighborhood'
maḥšar	*meḥşer*	*meḥşer*	'resurrection (day)'
ḥākim	*ḥakim*	*ḥakim*	'judge, governor'
ḥammām	*ḥemam*	*ḥemam*	'bath'
baḥr	*beḥr*	*beḥr*	'sea'

non-pharyngeal counterparts. Such is the word *haq* from Arabic *ḥaqq* 'right', or the Arabic word *ṭaʕm* 'taste' that is seen in eastern dialects of Northern Kurdish and in Central Kurdish without the voiced pharyngeal as *ṭam* and *tam*, respectively.

Furthermore, an original pharyngeal in a loanword may be substituted with the alternative pharyngeal sound, so, for example, the voiced pharyngeal of the Arabic *ṭamaʕ* 'greed' may be realized as either of the pharyngeals in different Kurdish dialects. Such indeterminate or alternative use of pharyngeals may exist within a single dialect (cf. Kahn 1976: 25). For instance, in the Mukri dialect of Central Kurdish, (Öpengin 2016: 41–42) the following Arabic-origin words can be found in both of the form pairs: *saʿib ~ saḥib* 'owner', *ʿerz ~ ḥerz* 'honour', *cemaʿet ~ cemaḥet* 'community'.

Finally, a pharyngeal may develop in loanwords that have no pharyngeal in the source language. Thus, in most of Northern Kurdish the Arabic word *ʔarḍ* 'earth' appears with a non-etymological pharyngeal as *ʿerd*, while the Arabic word *ǧāhil* 'naïve, young' is seen with a pharyngeal as *caḥêl* (but also *cahil*).

Although the pharyngeals in Kurdish occur mostly in Arabic loanwords, they have expanded also into inherited native Iranian lexicon, especially in Northern Kurdish. However, unlike in Arabic loanwords, fluctuation between pharyngeal and non-pharyngeal uses of such words among the dialects (sometimes in immediate geographic proximity) is readily apparent. Table 4 presents some native Iranian words of this kind. Where relevant, the non-pharyngeal forms are also noted, while Persian cognates are included for comparison.

More striking, however, is the emergence of a voiced pharyngeal in a subset of

Table 4: Pharyngeal sounds in native Iranian lexical items

Persian	Northern Kurdish	Central Kurdish	Gloss
abr	*ʿewr*	*hewr*	'cloud'
zabān	*ʿezman ~ ziman*	*ziman*	'language'
āsemān	*ʿesman*	*asman ~ hasman*	'sky'
ḫošk	*ḥişk ~ hişk*	*wişk*	'dry, hard'
haft	*ḥeft ~ heft*	*ḥewt*	'seven'
hašt	*ḥeşt ~ heşt*	*heşt*	'eight'
bahašt	*biḥeşt ~ bihişt*	*beheşt*	'paradise'

words with similar structure in the northern dialects of Northern Kurdish that are geographically farthest from direct Arabic/Semitic contact but close to Caucasian languages which also possess pharyngeals. Thus, the native words such as *masî* 'fish', *çav* 'eye', *mar* 'snake' (in Central Kurdish and in central and southern dialects of Northern Kurdish) appear in the northern dialects of Northern Kurdish with a pharyngeal, as *meʿsî, çeʿv, meʿr*. These are obviously the result of language-internal processes, though nested in an initial introduction of the phonemes into the language via contact with either Arabic or Caucasian languages, or both.

As for their distribution, the pharyngeal phonemes are most robustly present in the central areas of the Northern and Central Kurdish speech zones. Their presence in Arabic loanwords is weakened towards the extreme northern and southern peripheries in heavy contact with Turkish and Persian (cf. Map 1.27 in the Manchester Kurdish Database, which illustrates such weakening of pharyngeals at the peripheries through the distribution of the Arabic loanword *ḥeywan* 'animal').[6]

We turn now to the series of emphatic (pharyngealized) obstruents *ṭ, ḍ* and *ṣ, ẓ*. Table 5 gives a list of Arabic loanwords in which the original emphatic consonant is retained in Kurdish.

In the deeper-layer loanwords, the Arabic interdental and voiced alveolar emphatics are merged into the voiced emphatic alveolar phoneme *ẓ* in Kurdish. But in present-day Iraqi and Syrian Kurdish speech, especially those speakers with formal education may also pronounce the interdental phoneme, especially in the case of nonce borrowings and code-mixing.

On the other hand, quite a number of Arabic loanwords are pronounced with-

[6]See http://kurdish.humanities.manchester.ac.uk/pharyngeal-retentionloss-animal/.

Table 5: Arabic loanwords with emphatic consonants in Kurdish

Arabic	Northern/Central Kurdish	Gloss
ṭaʃm	*ṭam ~ṭeʿm/tam*	‘taste’
ṭāʔir	*ṭeyr*	‘bird’
baṭṭāl	*beṭal*	‘empty, cancel’
ðulm	*ẓulm*	‘oppression’
ḍābit	*ẓabit*	‘clerk’
ṣūfī	*ṣofî*	‘devotee, Sufi’
ṣāfī	*ṣaf/ṣafî*	‘clear’

out their original emphatic consonants, reanalysed thus as the corresponding plain phonemes (similarly to Persian), as in the items in Table 6.

Table 6: Arabic loanwords with lost emphatics in Kurdish

Arabic	Northern Kurdish	Central Kurdish	Gloss
ḥāṭir	*xatir*	*xatir*	‘mind’
ṭaraf	*teref*	*teref*	‘side, direction’
šayṭān	*şeytan*	*şeytan*	‘devil’
ḍaʃīf	*zeʿîf*	*zeʿîf*	‘weak’
ḥāḍir	*ḥazir*	*ḥazir*	‘ready’
qaṣṣāb	*qesab*	*qesab*	‘butcher’
fasīḥ	*fesîḥ*	*fesîḥ*	‘clear’
ṣabr	*sebr~ṣebr*	*sebr*	‘patience’

On the reverse side, some Arabic loanwords with no original emphatic consonants are pronounced with emphatic consonants in Kurdish, such as *ẓełał* (~ *ẓelal* and *zelal*) from Arabic *zulāl* ‘clear’ (dialectal *zalāl*), or *ẓelam* ‘man’ from Syrian Arabic *zalame*.

Finally, just as with the pharyngeal consonants, emphatic sounds also appear in inherited native Iranian words, as illustrated in Table 7.

Of the emphatic obstruents, the fricative pair (*ṣ, ẓ*) are found both in Northern and Central Kurdish, while the stops (*ṭ, ḍ*) are found only in Northern Kurdish, with the voiced counterpart being extremely rare. The fact that the voiceless emphatic stop is widespread only in Northern Kurdish most probably has to

Table 7: Emphatic consonants in native Iranian lexical items

Northern Kurdish	Central Kurdish	Gloss
meẓin	-	'big'
ẓiman	*ẓiman*	'language'
ẓik ~ zik	*zik*	'stomach'
aẓad	*aẓad*	'free'
ẓava	*ẓawa*	'groom'
beẓîn	*beẓîn*	'to run'
peẓ	*pez*	'sheep'
ṣal	*ṣał*	'year'
ṣed ~ sed	*ṣed*	'hundred'
ṣe	*ṣeg*	'dog'
beṣ ~ bes	*beṣ ~ bes*	'enough'
ṣawa	*sawa*	'very young, newborn'
ṣotin	*sûtan*	'to burn'
ṣiṣt	*ṣiṣt*	'loose'
ṭarî	*tarîk*	'dark'
ṭezî	*tezî*	'cold'
ṭeng	*teng*	'narrow'
ṭerm	*term*	'dead body'
ṭirş	*tirş*	'sour'
ḍaṣ(ik)	*das*	'sickle'
ḍirî	-	'blackberry bush'

do with the presence of two series of aspirated and unaspirated voiceless stops in the language (cf. Table 2). The unaspirated stops are probably intermediary in the development of emphatics. This is further reinforced by the fact that in Northern Kurdish the bilabial voiceless stop *p* also has an emphatic version, as in the native words *ṗeẓ* 'sheep' and *ṗenîr* 'cheese' (in some dialects; cf. Kahn 1976: 27). Within Northern Kurdish, they are found in more southerly dialects, and are noted to be particularly frequent in both the Kurdish and Neo-Aramaic of Duhok and Hakkari provinces (Blau 1989: 329). They tend to be less present moving northwards (Erzurum–Kars) while MacKenzie (1961: 43) notes that they are altogether absent in the Yerevan dialect. This distribution is of course consistent with a language-contact scenario, in the sense that in the northern dialects away from Semitic influence the language either did not develop emphatics or

lost them as a result of contact with and bilingualism in Armenian, Turkic and Caucasian languages that do not possess such emphatics.

Given the shallow history of written Kurdish, it is not possible to determine the historical period of the introduction of the emphatics and pharyngeals into the language. However, they are found abundantly even in the earliest Kurdish texts, especially in the Arabic component, but also in inherited lexical items, such as *ṣal* 'year', *ṣar* 'cold', *ṣed* 'hundred', *meẓin* 'big', *ḥemyan* 'all of them' (items taken from *Şêxê Senʿaniyan* by the early seventeenth-century poet Feqiyê Teyran, cf. Teyran 2011).

Three studies have treated the pharyngeals and emphatics in Kurdish, namely Kahn (1976), Anonby (forthcoming) and Barry (forthcoming). Barry (forthcoming) suggests that the pharyngeal sounds (including emphatics) in Kurdish are the result of contact influence from Arabic with a phonetic basis. The phonetic basis consists in the recategorization of vowels and the *h* sound within syllables with "flat" consonants (including pharyngeals, rhotics, grooved fricatives, and labials). Thus, initially, through extensive language contact with and bilingualism in Arabic, the speakers attained an active category of pharyngeals. Then the (inherited or loan) vocabulary with sounds that have pharyngeal-like effects on neighbouring vowels led to the reanalysis of the given vocabulary items as pharyngeal. In this account, the whole syllable is pharyngeal rather than individual sound segments. This account is particularly appropriate since, while it acknowledges the role of language contact with Arabic in the initial stage, it posits a phonetic mechanism of language-internal development of pharyngealization that captures an expansion of pharyngeals into historically non-pharyngeal lexical items that would be impossible to explain on purely language-contact grounds. It is, for instance, consistent with the fact that, in the above-presented data, the emphatics, but not pharyngeals in loan words, are restricted to the environment of more open vowels: *e*, *a*, *o*, and *i* [ɪ]. Furthermore, although not stated in the source study, the assumed subsequent development of a phonetic basis for the propagation of the pharyngeals into items originally without pharyngeal sounds is consonant with the facts of different stages or layers of borrowing. For instance, from the Arabic root *√ǧmʕ* we have three forms in Kurmanji: *civat* 'community, company', *cimat* 'the assembly of prayers in a funeral', and *cemaʕet* 'community'. The first form is probably the result of an early borrowing right after the initial Islamicization of the Kurds, as the fricativization of the bilabial nasal was active then (as seen also in *silav* 'greeting' from Arabic *salām*; Paul 2008). The second form with a slightly specialized semantic difference may have originated in a dialect where the mentioned fricativization did not occur. In any case, the first two

forms, which are clearly early borrowings, did not retain the original pharyngeal, whereas in a later borrowing from the same root, when one can assume that the pharyngeals were better tolerated in the language (and that the fricativization of bilabial nasal was not active), the pharyngeal sound did survive.

However, this account fails to explain why, in the great majority of the vocabulary with the relevant phonetic environment (syllables with "flat" consonants and low and back vowels), pharyngealization has not occurred. If the phonetic mechanism is integrated into the phonological system of the language, then pharyngealization would be expected in all relevant contexts. In this sense, although there is a phonetic ground to the propagation of the pharyngeals and emphatics in Kurdish, it may be safer not to postulate it as integrated into the phonological system of the language. Rather, the pharyngeals and emphatics should still be considered as peripheral to the phonological system (cf. Haig 2007; Anonby forthcoming), since, as noted by Haig (2007: 167), they are restricted to individual lexical items, their functional load is very limited, and there is considerable cross-speaker and cross-dialectal variability in the extent of their presence.

Although it is not the main focus of this chapter, a note on the reverse direction of contact influence is in order at this point. The Arabic dialects of Anatolia or Upper Mesopotamia (Mardin, Siirt, Kozluk, Sason, and the plain of Muş) have adopted some consonant and vowel phonemes via loanwords from Kurdish and Turkish, which do not exist in mainstream Arabic dialects (Jastrow 2011: 84; Akkuş, this volume: §3.1.1). The phonemes and example words with their sources are given in Table 8.

These additions into the phoneme inventory of the Anatolian Arabic are evidently the result of contact with Kurdish and Turkish. The introduction of these new phonemes has, as noted by Jastrow (2011: 84), on the one hand re-established the lacking symmetry caused by historical sound changes in Old Arabic, while on the other hand causing further sound shifts in the inherited Arabic vocabulary.

3.2 Morphology

It is usually assumed that Arabic influence on Kurdish is absent in the grammar (e.g. Edwards 1851), being largely restricted to phonology and lexicon. This is indeed to a large extent true. There are, however, several potential grammatical features that may be related to such contact influence.

Matras (2010: 75) suggests that the presence of aspect–mood prefixes in the languages of the Eastern Anatolian linguistic zone, namely Persian, Kurdish, Neo-Aramaic, Arabic and Western Armenian, is an outcome of language contact. Accordingly, all of these languages have a progressive–indicative aspectual

Table 8: Borrowed phonemes in Arabic dialects of Anatolia

Phoneme	Example
bilabial stop /p/	*parčāye* 'piece' < Tr. *parça*
voiced labio-dental fricative /v/	*davare* 'ramp' < Kr. *dever* (F) 'place'
voiceless affricate /č/	*čəqmāq* 'lighter' < Tr. *çakmak*[a]
voiced palatal fricative /ž/[b]	*ṭāžī* 'greyhound' < Kr. *ṭajî*
voiced velar stop /g/	*gōmlak* 'shirt (modern)' < Tr. *gömlek*
mid long front vowel /ē/	*tēl* 'wire' < Tr. *tel* (via Kr. *têl*)
mid long back vowel /ō/[c]	*ḫōrt* 'young man' < Kr. *xort*

[a]It is more probable that this word (and others attributed to Turkish) is borrowed via Kurdish, since the uvularization (/k/ > /q/) in loanwords and the change in the vowel of the first syllable (cf. also *qeymaẍ* 'cream', from Turkish *kaymak*) are typical of Kurmanji spoken in the region.

[b]Note that the reflex of Arabic ⟨ج̲⟩ in this variety is /ǧ/, not /ž/.

[c]Note also that the original Arabic diphthongs *ay and *aw are preserved in this variety, not monophthongized to /ē/ and /ō/.

prefix (in turn: *mī-, di-, gə-, ko-, ba-/-a-*), while subjunctive is marked either by the absence of the indicative prefix (Armenian and Neo-Aramaic) or by a specialized subjunctive prefix (Persian, Kurdish, Arabic). Since such aspect–mood prefixes are considered typical of Iranian languages of the region, they would have diffused from Kurdish and Persian into the other languages of the zone, including Arabic (which in its standard grammar does not have such forms; Ryding 2014: 46–47). However, assessing the validity of Eastern Anatolia's being a linguistic area, Haig (2014: 20–25) casts doubt on this claimed contact scenario, primarily since: (i) the feature exists in Arabic dialects outside the region; and (ii) it is absent in the two major languages of Anatolia, namely Turkish and Zazaki. Jastrow (2011: 92), on the other hand, although acknowledging the source of such verbal prefixes grammaticalizing from Old Arabic verb forms, hypothesizes – though without providing supporting arguments – that they may have developed under Turkish and Kurdish influence. Assessing also the grammaticalization of such formatives in various languages and rejecting a contact scenario behind their frequent occurrence in the languages of Anatolia, Haig (2014: 26) concludes that the present indicative prefixes found in Kurmanji, and in certain varieties of Aramaic and Arabic in Anatolia, could be interpreted as reflexes of an inherited morphological template, which is well-attested in the related Northwest Iranian and Semitic languages outside Anatolia.

Another (not previously discussed) candidate for Arabic influence on Kurmanji Kurdish relates to gender assignment in more recent loanwords from European languages. In Kurmanji, like Arabic, nouns are assigned to feminine and masculine genders. The gender of inanimate nouns is largely arbitrary, with limited morpho-phonological basis in both languages. In Arabic, words carrying the *-a* ending are feminine, while in Kurmanji abstract nouns ending in *-î* are feminine, while the rest may be of either gender. Now, when Arabic borrows modern vocabulary items from European languages, items ending in *-a* are assigned to feminine gender, while the rest are assigned to masculine gender (Ibrahim 2015: 5). The default gender assigned to new lexical borrowings is masculine in Arabic. There is as yet no research on the gender assignment of borrowings in Kurmanji. However, it is easily observed that Kurmanji spoken in Turkey mostly favors feminine, while the Kurmanji of Iraq uses masculine gender for integrating modern vocabulary items into the language. The modern lexical borrowings (boldface) in (1) are all assigned to masculine gender in Badini Kurmanji of Iraq. Note that the gender of the nouns is visible in the *ezāfe* (see §3.3) and oblique case suffixes.

(1) Badini dialect of Kurmanji in Iraq (from media outlets)

a. **sîstem**-ê endroyd-ê
system-EZ.M android-OBL.F
'Android system'

b. serok-ê **parlemanî**
president-EZ.M parliament-OBL.M
'the president of the parliament'

c. **form**ê têgehişt**in**-ê
form-EZ.M understanding-OBL.F
'the form of understanding'

d. **moral**-ê diyalog
moral-EZ.M dialogue
'the moral of dialogue'

e. **proj(e)**-ê av-ê
project-EZ.M water-OBL.F
'the water project'

f. **prensîp**-ê hevwelatîbûn-ê
principle-EZ.M citizenship-OBL.M
'the principle of citizenship'

Now, all of these lexical borrowings exist also in Kurmanji as spoken in Turkey, but they are systematically used with feminine gender. For instance the phrase in (1b) would be realized as *serok-ê parleman-ê* (president-EZ.M parliament-OBL.F), with the feminine form of the oblique case suffix.

As was stated above, the majority of such modern lexical borrowings in Arabic are assigned to masculine gender. The masculine gender assignment in Kurmanji in Iraq is thus most probably motivated by the Arabic gender assignment pattern. This is all the more plausible when we consider that Arabic, as the dominant state language for the Iraqi Kurds for almost a century, serves also as the intermediary language via which such lexical items are normally borrowed into Kurmanji in Iraq. However, this contact influence must have been established relatively recently, since earlier technical borrowings in Kurmanji in Iraq such as *têlevizyon* and *radyo* are treated as feminine nouns, despite being masculine in Arabic.

3.3 Syntax

Although several studies have dealt with the outcomes of language contact between Kurdish and (Neo-)Aramaic in the grammar of these languages – especially on such topics as alignment (Coghill 2016), word order (Haig 2014), and noun phrase morphology (Noorlander 2014) – as far as I am aware, the only study on Arabic–Kurdish contact in grammar is the short note of Tsabolov (1994) about the distinctive position of the possessor in a multiple-modifier noun phrase in Northern Kurdish.

As is well known, a number of West Iranian languages (Middle and contemporary Persian, Kurdish, Zazaki, etc.) employ a vocalic bound morpheme for linking post-head modifiers in a noun phrase, called *ezāfe* or *izāfe* (from Arabic *ʔiḍāfa* 'joining, addition'), as in (2) and (3).

(2) Persian (personal knowledge)
ẖāna-e bozorg
house-EZ big
'(the) big house'

(3) Northern Kurdish (personal knowledge)
xanî-yê mezin
house-EZ.M big
'the big house'

The *ezāfe* in Northern Kurdish differs from its cognates in, for instance, Central Kurdish and Persian, as it inflects for gender (masculine *-ê* vs. feminine *-a*) and number (singular *-ê/-a* and plural *-ên/-êd*), in addition to having secondary or pronominal forms used in chain *ezāfe* constructions with multiple modifiers (and some other predicative functions; cf. Haig 2011; Haig & Öpengin 2018). In most West Iranian languages, noun phrases with multiple modifiers have their head noun first, followed by qualitative then possessive modifiers, as in (4) and (5). This is also the order in Middle Persian, as in (6), where Tsabolov (1994: 122) considers such constructions may be regarded as prototypes of the *ezāfe* constructions of modern West Iranian languages.

(4) Persian (personal knowledge)
ẖāna-e bozorg-e Malek
house-EZ big-EZ PN
'Malek's big house'

(5) Central Kurdish (personal knowledge)
kurr-î gewre-y Karwan
son-EZ big-EZ PN
'my friend's beautiful daughter'

(6) Middle Persian (Tsabolov 1994: 122)
pus ī mas ī Artavān
son EZ big EZ PN
'Artavan's elder son'

However, in Northern Kurdish the order of modifiers is reversed, such that a possessor of the head noun in the noun phrase comes before attributive modifiers, as in (7), where the secondary linking element is glossed as SEC.

(7) Northern Kurdish (personal knowledge)
xanî-yê Malik-î (y)ê mezin
house-EZ.M PN-OBL.M EZ.M.SEC big
'Malik's big house'

Tsabolov observes that these syntactic particularities of Northern Kurdish have no parallels in other Kurdish varieties and Iranian languages as a whole, but that they correspond to the word order in noun phrases in Arabic, as can be seen in the comparison of (8) and (9).

(8) Arabic (Tsabolov 1994: 123)
miḥfaðatu ṭ-ṭālibi l-ğadīdatu
bag.NOM DEF-student.GEN DEF-new.F.NOM
'the student's new bag'

(9) Northern Kurdish (Tsabolov 1994: 123)
çent-ê şagirt-î taze
bag-EZ.M student-EZ.M.SEC new
'the student's new bag'

Note that although in standard Kurmanji (Northern Kurdish) the forms of the primary and secondary *ezāfes* are identical, with the difference being in the latters' status either as enclitics or independent particles, in the northern dialect of Northern Kurdish considered by Tsabolov, the singular forms of the secondary *ezāfe* are different (with masculine *-î* and feminine *-e*). In Tsabolov's view, the centuries-old close contacts between Kurdish and Semitic dialects, especially Arabic, have not only resulted in the above-described change of noun-phrase-internal word order (syntactic) but also in the development of secondary forms of *ezāfe* through the "weakening" of the primary ones (morphological), because, he argues, such distinct forms "were necessary for correlating each attribute in an [*ezāfe*] chain with the ruling noun they refer to" (1994: 123).

On closer scrutiny, however, the motivation Tsabolov puts forward for the morphological change may not be entirely correct, since, on the one hand, *ezāfe* forms in Northern Kurdish distinguish gender and number, which already correlate the modifiers with their head nouns, and on the other hand, in the majority of Northern Kurdish dialects the primary and secondary *ezāfes* are formally identical. The change in form is an instance of vowel raising (*a* > *e*, *ê* > *î*) that is also observed elsewhere in the morphology of noun phrase (cf. Haig & Öpengin 2018).

As for Tsabolov's main claim regarding word-order change leading to the initial positioning of a possessor modifier in the noun phrase, here too the role of language contact might require revision, since it might have more to do with language-internal organization of morphological material: Zazaki (geographically contiguous with Kurmanji but from a separate historical source to Kurdish), which, like Kurmanji, has gender/number-marking *ezāfe* forms and a case distinction in its nominal system, follows precisely the same word order pattern as Kurmanji in the noun phrase (cf. Todd 2002: 95), while Sorani, which has lost gender/number-marking in *ezāfes* and case distinctions in its nominal system, differs from them following the Persian and Middle Persian pattern (cf.

Öpengin 2016: 61–64). That is, the determining factor seems to be the presence or absence of gender/number-marking *ezāfe* forms, which enable reference tracking between heads and dependents in a noun phrase independently of word order.

Despite the scepticism one may have towards Tsabolov's hyopthesis, there is a rather parallel more recent syntactic change in progress stemming from the Arabic influence on the Kurdish of Iraq. This change concerns especially the naming of institutions, such as schools and airports. Recall that in Central Kurdish the possessor in a chain *ezāfe* construction is positioned at the end of the noun phrase, as illustrated in (5). However, in the case of these examples, the proper name occurs right after the head noun and before the qualitative modifier, as in (10) and (11).

(10) Central Kurdish (official signage)
qutabxane-y Qemeryan-î seretayî
school-EZ PN-EZ primary
'Qamaryan primary school'

(11) Central Kurdish (official signage)
firokexane-y Hewlêr-î nêwdewletî
airport-EZ PN-EZ international
'Hawler international airport'

If the proper name is understood as having the function of possessor here, this is an order that is rather different from the typical Central Kurdish syntax of chain *ezāfe* constructions. But this is precisely the order described for multiple modifier noun phrases of Arabic, as in (8). Thus the order in (11) is the exact replication of the Arabic version of the same name illustrated in (12).

(12) Arabic (official signage)
maṭār arbīl ad-dawlī
airport PN DEF-international
'Erbil international airport'

This is clearly a recent imposition from Arabic which does not seem to have gone much beyond naming institutions, especially official signage: the Arabic-like ordering of the name of the airport appears only half as frequently as the inherited order in a Google search. Furthermore, there is no trace of such a word order pattern in the use of Central Kurdish in Iran.

3.4 Lexicon

Arabic influence on Kurdish and all other Near Eastern languages is observed most clearly and abundantly in the vast number of loanwords. According to Perry (2005: 97), the process of lexical convergence initially took place in Persian between the ninth and thirteenth centuries, when a large number of learned terms were borrowed into literary Persian, and thence transmitted to the other languages of the region. This scenario explains some of the similarities of loanword integration in the two languages (e.g. the borrowing of *tāʔ marbūṭa* as *-at/-et* (rather than *a*) in a number of words, such as *hukūmat* 'government', Persian *hokūmat*, and *quwet* 'strength', Persian *qovvat*). However, being spoken in a region that is closer geographically to Arabic-speaking communities, and having had its own educational and religious institutions where Arabic served as the high literary language, Kurdish must have also followed its own course of contact with Arabic. Despite this, there are no studies of lexical borrowing from Arabic into Kurdish. Given the vastness of the topic, with its layers of time-depth and subsantial extra-linguistic aspects, I can only propose here to sketch the major lexical domains of borrowing, and note some observations on the word class and morpho-phonological integration of the borrowings.

The three major varieties differ in their proportions of borrowing from Arabic. Impressionistically, Northern Kurdish seems to have borrowed most extensively. There is, however, a deeper layer of lexical borrowings shared throughout Kurdish (some of which are common to all or most of the Near Eastern languages), such as the following (cited in their Northern Kurdish forms):[7]

(13)

xerab 'bad'	< Ar. *ḫarāb* 'ruins'
xelk/xelq 'people'	< Ar. *ḫalq* (*√xlq* 'to create')
xiyanet 'betrayal'	< Ar. *ḫiyāna*
xizêm 'nose-ring'	< Ar. *ḫizām*
xizmet 'service'	< Ar. *ḫidma*
ʿeql/aqil 'reason'	< Ar. *ʕaql* (*qəltu* Ar. *ʕaqəl*)
qelem 'pen'	< Ar. *qalam*
quwet 'strength'	< Ar. *quwwa*
kitêb 'book'	< Ar. *kitāb*
xiyal 'thought, grief'	< Ar. *ḫayāl* 'imagination'
hevîr 'dough'	< Ar. *ḫamīr*

[7]The main source for the lexical items in this section, together with the information regarding their Arabic origin, is Chyet (2003). However, I have supplied the interpretation and the discussion of the material and as such only I am responsible for any shortcomings.

fikr 'thought, idea'	< Ar. *fikr*
fêkî/fêqî 'fruit'	< Ar. *fākiḥa*
ḥal 'condition'	< Ar. *ḥāl*
ḥazir 'ready'	< Ar. *ḥāḍir*
şol/şuẍul 'work'	< Ar. *šuɣl*
terk 'abandonment'	< Ar. *√trk* 'to abandon'

Within varieties too, the dialect zones where the communities have had historically closer contact with Arabic-speaking areas show greater Arabic influence in vocabulary. Thus, the dialect of Northern Kurdish named as Southern Kurmanji by Öpengin & Haig (2014), spoken around Mardin and Diyarbekir provinces in Turkey, the Jazira province of northeast Syria, and the Sinjar region of Iraq, is the dialect with most extensive Arabic lexical borrowings. Thus, the following items are restricted to this dialect of Northern Kurdish: *tefa-ndin* 'extinguish-TR.INF' (from dialectal Ar. *ṭafa* or standard *ṭafiʔa*), *şiteẍl-în* 'speak-INTR.INF' (from dialectal Ar. *ištaɣal* 'to work'), *hersim* 'unripe and sour grapes' (from Ar. *ḥiṣrim*), *siʿûd* 'good luck' (Ar. *suʕūd*, pl. of *saʕd*), *şîret* and *şêwr* 'advice, counsel' (Ar. *√šwr*).

Arabic loanwords in Kurdish belong to various semantic fields, such as kinship, body parts, animals, agriculture, basic tools, temporal concepts and religion. Regarding kinship terms, while the terms for the members of the nuclear family are all inherited, the four second-degree kin terms are all borrowed from Arabic: *met* 'paternal aunt' (cf. Ar. *ʕamma(t)*; this item does not exist in Sorani), *xalet/xaltî* 'maternal aunt' (Ar. *ḫāla*), *mam ~ am* 'paternal uncle' (Ar. *ʕamm*), *xal* 'maternal uncle' (Ar. *ḫāl*). Considering that the language had its own kin terms before its contact with Arabic, the borrowing of such kin terms constitutes a case of prestige borrowing, probably motivated by the use of such kin words as address forms in the cultures of the region (cf. Haig & Öpengin 2015).

Similarly, while words for basic animals are inherited, the animals not indigenous to the mountainous region of core Kurdistan are borrowed from Arabic, such as *tîmseḥ* 'crocodile' (Ar. *timsāḥ*), *fîl* 'elephant' (Ar. *fīl*), *xezal* 'gazelle, deer' (Ar. *ɣazāl*). Likewise, the generic term for 'bird' or 'large birds' is the Arabic loanword *ṭeyr* (Ar. *ṭayr*), while the category word *ferx* 'young bird/chicken' is also from Arabic *farḫ*. Several agricultural terms are also borrowed from Arabic, such as *ẓad* 'grain, food' (Ar. *zād* 'provisions'), *simbil* 'spike (of corn or wheat)' (Ar. *sunbul*), *xox* 'peach' (Ar. *ḫawḫ*), *dims* 'grape molasses' (Ar. *dibs*). Various terms for spaces and tools of daily life are also borrowed from Arabic, such as *saʿet* 'hour' (Ar. *sāʕa*), *sifre* 'tablecloth' (Ar. *sufra* 'dining table'), *qefes* 'cage' (Ar. *qafaṣ*), *ḫubr* 'ink' (Ar. *ḥibr*), *ḥemam* 'bath' (Ar. *ḥammām*), *ḥewş* 'yard' (Ar. *ḥawš*), *meẍmer* 'velvet' (Ar. *muḫmal*). Some occupational terms from Ara-

bic are *neqş* 'embroidery' (Ar. *naqš* 'painting, drawing'), *ḥedad* 'blacksmith' (Ar. *ḥaddād*), *ʿesker* 'soldier' (Ar. *ʕaskar* 'army'), *tucar* and its older form *têcirvan* (Ar. *tuǧǧār* 'traders', sg. *tāǧir*).

The older layer of administrative and legal terms are predominantly derived from Arabic – though they may have mostly entered via Persian and Ottoman Turkish – such as *sultan* 'monarch' (Ar. *sulṭān*), *walî* 'provincial governor' (Ar. *wālī*), *muxtar* 'village chief' (Ar. *muḫtār*), *ḥukûmet* 'government' (Ar. *ḥukūma*), *meḥkeme* 'court' (Ar. *maḥkama*), *deʿwā* 'request, court case' (Ar. *daʃwa* 'request, invitation' and *daʃwā* 'court case'), *qanûn* 'law' (Ar. *qānūn*), *mekteb* 'school' (Ar. *maktab* 'office, desk').

As for religious terms, similar to the Persian case (cf. Perry 2012: 72), a good number of basic Islamic concepts are inherited, such as the words for god, prophet, angel, devil, heaven, purgatory, prayer, fasting, and sin. In some instances, the Arabic equivalents of these terms exist alongside the inherited ones, restricting the use of the latter, as in the cases of *şeytan* 'devil' and *cehnem* 'hell', from Arabic *šayṭān* and *ǧahannam*, replacing the Iranian *dêw* and *dojeh*. Many other basic and more peripheral concepts are borrowed from Arabic, such as the following: *xêr* 'good' (Ar. *ḫayr*), *xezeb* 'wrath' (Ar. *ɣaḍab*), *civat/cemaʿet* 'society, gathering' (Ar. *ǧamāʿa*), *ḥec* 'pilgrimage' (Ar. *ḥaǧǧ*), *şeytan* 'devil' (Ar. *šayṭān*), *weʿz* '(Islamic) sermon' (Ar. *waʿẓ*), *ḥelal* 'permitted' (Ar. *ḥalāl*), *ḥeram* 'forbidden' (Ar. *ḥarām*), *ruḥ* 'soul, spirit' (Ar. *rūh*), *tizbî* (Sorani *tezbêḫ*) 'prayer beads' (Ar. *tasbīḥ*).

Finally, there are also a large number of concepts (temporal, moral, cosmological) that originate from Arabic roots, such as *sibe(h)* 'morning, tomorrow' (Ar. *ṣabāḥ*), *heyam* 'period' (Ar. *ayyām* 'days'), *hêsîr* 'prisoner' (Ar. *ʔasīr*), *dinya* 'world' (Ar. *dunyā*), *ḥesab* 'count, calculation' (Ar. *ḥisāb*), *ḫîle* 'trick, ruse' (Ar. *ḥīla*), *ḥel* 'solution' (Ar. *ḥall*), *eşq* 'love' (Ar. *ʕišq*), *ʿerz* 'honor, esteem' (Ar. *ʕirḍ*). Note also that the word *dinya* is used corresponding to the English expletive subject *it* in time and weather expressions, as in *dinya esr e* 'it is late afternoon' or *dinya ewr e* 'it is cloudy'. This usage is noted to exist also in colloquial Arabic (Chyet 2003: 155).

Some other interesting developments with Arabic material in Kurdish lexicon may be noted here. The Arabic *daʃwa* 'invitation' has resulted in two related but different concepts: *dawet* 'wedding ceremony' and *deʿwet* 'invitation'. While the latter meaning is shared in Ottoman/Turkish and Persian, the former is a Kurdish-internal semantic expansion of the source meaning. The Kurdish (in all three varieties) word for 'home' *mal*, in the sense of family and familial belongings, rather than the house as a structure, is probably derived from the Arabic word *māl* 'goods, property'. The generic term in Kurdish that designates Chris-

tians regardless of their ethnicity and confession is *fileh/file* which derives from Arabic *fallāḥ* 'peasant, farmer'. Finally, there is the word *mixaletî* 'the son of the maternal uncle or aunt' in the southern Kurmanji dialect of Northern Kurdish that can probably be analysed as *mi* (< *ben* 'son') + *xalet* 'aunt' (< Ar. *ḫāla*) + *î* 'my'.

Turning now to the word class categories of the loanwords, as has been seen from the presentation of semantic domains above, most Arabic loanwords in Kurdish are nouns. However, many Arabic noun loans are incorporated into Kurdish verb forms. This takes place either through morphological integration or syntactic composition. In morphological integration, the Arabic root (whether nominal or verbal) is taken as the stem onto which the Kurdish verbal suffixes *-în/-îyan* for instransitives and *-andin* for transitives are added. Thus the Arabic noun *ʕilm* 'knowledge', apart from being used in its nominal sense, serves as the stem for the derivation of the intransitive *ʿelimîn* (*ʿelim-în*) 'to learn' and transitive *ʿelimandin* 'to teach, educate'. The following verbs are further examples of using Arabic roots (whether the original borrowings are nouns or verbs is not always clear) in the derivation of verbs in Kurdish: *tefandin* 'to extinguish' (Ar. *ṭafa/ṭafiʔa*), *fetisandin* 'to suffocate' (Ar. *faṭṭas*), *fetilîn* 'to turn around' (?Ar. *fatala* 'to twist together'), *qulibîn* 'to be overturned' (Ar. *qalaba* 'to overturn'), *sekinîn* 'to stand, stop' (Ar. *√skn* 'calm, rest'), *fikirîn* 'to think; to look at' (Ar. *fikr* 'thought').[8] The verb *qelandin* 'to roast; to uproot' has two sources as Ar. *qalā* and *qalaʃa,* respectively, which explains its polysemy in Kurdish.

In syntactic composition, on the other hand, a compound verb[9] is obtained by combining an Arabic root with an inherited auxiliary light verb, such as *kirin* 'do' or *dan* 'give' for transitives, and *bûn* 'to be' for instansitives. Thus, the combination of Arabic adjective loanword *xerab* 'bad' (< Ar. *ḫarāb* 'ruin') with *kirin* yields the verbal meaning 'to destroy' while its combination with *bûn* means 'to go bad, be spoiled'. Some example compound verbs with Arabic roots are given in (14).

[8]Kurdish possesses a number of preverbs such as *ve-* and *ra-*. When inflected with tense–aspect–mood prefixes, these preverbs are detached from the verb stem, as with the verb *ve-kirin* 'to open' in *ve-di-ki-m* (PVB-IND-do.PRS-1SG) 'I open (it)'. Now, the initial syllable of the verbs *sekinîn* and *fekirîn*, which are based on Arabic loan roots, resemble such Kurdish preverbs. As a result, in some dialects, they are treated as preverbal elements detaching from the verb stem, as with *fe-di-ki-m-ê* 'I look at it' (own data, Şirnak area) or *se-di-kin-e* 's/he stands' (own data, from Gevaş), where the initial syllables of originally Arabic roots are reanalysed as preverbs.

[9]Here the term *compound verb* is employed in a pre-theoretical sense, regardless of whether or not the given complex verb is considered to form a compound. See Haig (2002) for a discussion of complex verbs in Kurdish.

(14) *qedr* 'respect' (Ar. *qadr*) + *girtin* 'to hold' = 'to respect'
silav 'greeting' (Ar. *salām* 'peace') + *dan* 'to give' = 'to greet'
teʿn/ṭan 'scolding' (Ar. *ṭaʃn* 'piercing') + *dan* = 'to criticize'
qedeӧe 'forbidden' (Ar. *qadaḥa* 'to rebuke') + *kirin* 'to do' = 'to forbid'
qesd 'intention' (Ar. *qaṣd*) + *kirin* = 'to head for'
zeʿîf 'weak' (Ar. *ḍaʃīf*) + *bûn* 'to be' = 'to become slim'

What motivates the choice between the morphological versus syntactic technique in the integration of Arabic loan roots in forming verbs in Kurdish is not yet clear. While a few such verbs are found to be used in both synthetic and analytic forms, such as *ceribandin* and *cerebe kirin* 'to try' (Ar. < *ǧarraba*), most verbs are used in just one of the two forms. However, there is a great deal of dialectal differentiation as to whether a verb is analytically or synthetically integrated. Thus, the morphologically integrated verbs of most Northern Kurmanji dialects such as *emilandin* 'to use' (dialectal Ar. *ʃimil* 'to do'), *şuӿulîn* (Ar. *šuɣl* 'work'), *fikirîn* (Ar. *fikr* 'thought') are seen in the southeastern Badini dialect in synthetic form, with a nominal base combining with a light verb, as *emel kirin*, *şol kirin*, *fikr kirin*.

There are also various function words (discourse markers, conjunctions, adverbs) which are either borrowed from Arabic or developed in Kurdish based on material borrowed from Arabic. Thus, the conjunction *xeyrî* (also seen as *xeyr ji* and *xêncî*) 'apart from, besides' is based on Arabic *ɣayr* 'other than', while the adversative *emā* 'but' is dervied from Arabic *ʔammā* 'however'. The similative *şibî* (also *şubhetî* and *şitî*) is derived from the Arabic root *√šbh* 'resemblance'. The classifiers *ḥeb* (and the adverbial *hebekî* 'a little') and *lib* are derived from Arabic *ḥabb* 'grain(s)' and *lubb* 'kernels', respectively. Finally, some discourse and verbal adverbs resulting from Arabic sources are as follows: *meselen* 'for instance' and *helbet* 'of course' are from Arabic *maθalan* and *al-batta*; in the eastern section of the Badini dialect of Kurmanji, there is the use of the discourse marker *seḥî* 'apparently, that means', which is derived from the Arabic *aṣaḥḥ* 'more correct' – which separately exists in wider Kurdish as *esseḥ* 'certainly'; while, finally, the Arabic adjective *qawī* 'strong' has evolved into an adverb *qewî* 'very; very much' (though this is more literary than spoken).

All of these lexical borrowings illustrate matter transfer (in the sense of Matras & Sakel 2007). In the following we have two instances of pattern transfer. First, there is a particular adverbial form *nema* 'no longer', found only in the Southeastern dialect of Kurmanji, spoken in the Mardin region of Turkey and Jazira region of northeast Syria. This can be analysed as *ne-ma*, consisting of the negative prefix *ne-* and the past tense 3SG conjugation of the verb *man* 'to stay', as

in (15).

(15) Southern dialect of Northern Kurdish (Media)[10]
nema di-kar-im veger-im welêt
no.longer IND-be.able.PRS-1SG return.PRS.SBJV-1SG country.OBL
'I can no longer return to the homeland'

There is an immediately corresponding adverbial form *mā ʕād* 'no longer' in Arabic, which is based on the negative form of the semantically similar verb *ʕād* 'to return, keep doing'. This is obviously not a very recent development as it is shared in the whole dialect area across country borders, but seemingly not so deep either as to be shared by all Kurdish varieties, not even by all Northern Kurdish dialects, further strengthening the particular status of the Jazira region in Arabic–Kurdish language contact.

Second, there is a particular lexical construction *bi X rabûn* 'to do; to complete; to achieve' in Northern Kurdish and *hellsan be X* in Central Kurdish, where *X* stands for any activity or task (usually in the form of an infinitive verb). The construction is based on the verb for 'to rise, stand' and a preposition in both varieties, as illustrated in (16) and (17).

(16) Central Kurdish (Media)[11]
polîs hellsa be kokirdinewe-y zanyarî
police rise.PST.3SG with collecting-EZ information
'The police undertook (the task of) collecting information.'

(17) Northern Kurdish (Media)[12]
Mîr Celadet (…) bi kar-ê dewlet-ek-ê rabû
Emir Celadet with work-EZ.M state-INDF-OBL.F rise.PST.3SG
'Emir Celadet undertook the work of a state.'

This lexical construction also has a parallel in Modern Standard Arabic, based on the verb *qāma* 'to stand (up)' and the preposition *bi* 'with', with the collocation *qāma bi* meaning 'to undertake'. This is obviously a recent influence on Kurdish, as it is seen only in Iraq and Syria, and in a manner cross-cutting the broad variety borders between Sorani and Kurmanji.

[10]From a poem by an author from Syria, available online at: http://avestakurd.net/blog/2016/10/26/romanivs-kurd-jan-dost-lal-b-ye-vdyo/

[11]URL of article: http://www.kurdistan24.net/so/news/5ca67132-7a7f-4840-bfb4-dea5bf25ea2e

[12]URL of article: http://portal.netewe.com/mir-celadet-bedirxan-bi-tene-sere-xwe-bi-kare-dewleteke-rabu/

4 Conclusion

Contact with Arabic, which started in the early medieval period (approx. 7th–8th centuries) with the arrival of Islam in the Near East, has had a profound impact on Kurdish, particularly on its lexicon and phonology. Given the total absence of any substantial previous study on the matter, the present chapter provides a first assessment of the influence of Arabic on Kurdish, primarily as represented in Kurdish phonology and lexicon but also, albeit more restrictedly, in morphology and syntax. Kurmanji Kurdish seems to be the variety that is most affected by contact with Arabic, which is understandable considering the geographical continuity of the Kurdish and Arabic communities, especially in the historical Jazira region and more widely in Upper Mesopotamia (in Mardin–Diyarbekir, Mosul–Sinjar, and Haseke province). There are thus areas which show more intensive Arabic influence within the speech zones of major Kurdish varieties, while the outcomes of the contact reflect different layers in terms of time depth. Accordingly, the deeper-layer influence comes in the form of lexical convergence with Arabic, sometimes through the intermediary of Persian and/or Ottoman Turkish. This contact has repercussions in the expansion of the phonological inventory of the language, and is shared across most Kurdish varieties. There are no unquestionably demonstrated changes in the morphosyntax resulting from contact with Arabic at this layer. At the relatively shallower layer, the influence is mainly seen in Syria and Iraq, and in the form of further expansion of the phonological inventory and a vocabulary heavily lexified by Arabic roots incorporated also into the verbal domain. There are also several cases illustrating morphosyntactic and lexicosyntactic change, such as the default gender assignment and word order in complex noun phrases, as well as certain phrasal and adverbial lexical items.

In terms of "cognitive dominance", in the sense of Van Coetsem (1988; 2000) and Lucas (2015), in these instances of contact influence, the deeper-layer influence, which is restricted to, or related to, lexical borrowing, takes place with the speakers being cognitively dominant in the recipient language, Kurdish. The more recent instances of heavy lexification, and morphosyntactic and lexicosyntactic changes may, however, be the result of imposition, where the speakers are dominant in the source language.

These outcomes may also be related to bilingualism and language configuration in historical perspective. That is, the absence of imposition (in the form of morphosyntactic changes) in the deeper historical layer, and the restriction of the influence to lexicon, point to the absence of widespread Arabic–Kurdish bilingualism among the speakers of Kurdish at those historical stages. Some im-

position of this kind is observed in the Kurmanji of the Jazira region, which is known to have had greatest speaker contact between Kurdish and Arabic speech communities. By contrast, the widespread bilingualism and Arabic-dominant linguistic configuration in Syria and Iraq for at least a century has led to instances of imposition where the morphosyntactic and lexical patterns of Arabic are replicated in Kurdish. These outcomes are also mostly consonant with the predictions of Van Coetsem's (1988; 2000) "stability gradient", which argues that lexicon is less stable than syntax and phonology, which require dominance in the source language in order to be affected by contact-induced change.

Given the limitations of a first attempt, much is yet to be explored regarding Kurdish–Arabic language contact. In particular, the precise paths of development of pharyngeals and emphatics in Kurdish should be analysed through fieldwork-based comparative dialect data, while, in the domain of lexicon, it is important to analyse the morphophonological integration of borrowings into Kurdish. It is also of interest to be able to develop diagnostics to disentangle direct Arabic influence on Kurdish from influence via other major languages such as Persian and Ottoman Turkish. Finally, a detailed account of the history of Kurdish–Arabic socio-political and cultural contact is of utmost importance in order to complement the linguistic data and enable a more fine-grained analysis of the agentivity of contact-induced change in Kurdish.

Further reading

Barry (forthcoming) is a comprehensive and theoretically grounded treatment of the introduction and further propagation of pharyngeal sounds in Kurdish.
Chyet (2003) is the most comprehensive Kurdish–English dictionary, providing information on the source language of most loanwords in Kurdish, including those from Arabic.
Tsabolov (1994) is the only work published so far on Arabic influence on the grammar of Kurdish.

Acknowledgements

I would like to thank the editors of the volume and an anonymous reviewer for their helpful and detailed feedback. Thanks also to Adam Benkato for his help with Arabic data. Only I am responsible for any remaining shortcomings and errors.

Abbreviations

Ar.	Arabic
BCE	before Common Era
ca.	circa
DEF	definite
DRCT	directional
dial.	dialectal
EZ	*ezāfe*
F	feminine
GEN	genitive
Kr.	Kurdish
IND	indicative
INDF	indefinite
INF	infinitive
INTR	intransitive
IPFV	imperfective
M	masculine
NEG	negative
NOM	nominative
OBL	oblique
PL/pl.	plural
PN	proper noun
POSS	possessive
PRS	present
PST	past
PVB	preverb
RL	recipient language
SBJV	subjunctive
SEC	secondary or pronominal *ezāfe*/linking element
SG/sg.	singular
SL	source language
TR	transitive
Tr.	Turkish

References

Anonby, Erik. Forthcoming. Emphatic consonants beyond Arabic: The emergence and proliferation of uvular–pharyngeal emphasis in Kumzari. *Linguistics*.

Barry, Daniel. Forthcoming. Pharyngeals in Kurmanji Kurdish: A reanalysis of their source and status. In Songül Gündogdu, Ergin Öpengin, Geoffrey L. J. Haig & Erik Anonby (eds.), *Current topics in Kurdish linguistics*. Bamberg: University of Bamberg Press.

Blau, Joyce. 1989. Le kurde. In Rüdiger Schmitt (ed.), *Compendium linguarum iranicarum*, 327–335. Wiesbaden: Reichert.

Bois, Thomas, Vladimir Minorsky & David N. MacKenzie. 2012. Kurds, Kurdistān. In Peri Bearman, Thierry Bianquis, C. Edmund Bosworth, Emeri Johannes van Donzel & Wolfhart P. Heinrichs (eds.), *Encyclopaedia of Islam*, 2nd edn (online). Leiden: Brill.

Chyet, Michael L. 2003. *Kurdish–English dictionary*. New Haven, CT: Yale University Press.

Coghill, Eleanor. 2016. *The rise and fall of ergativity in Aramaic: Cycles of alignment change*. Oxford: Oxford University Press.

Edwards, Bela B. 1851. Note on the Kurdish language. *Journal of the American Oriental Society* 2. 120–123.

Folmer, Margaretha L. 1995. *The Aramaic language in the Achaemenid period: A linguistic study in variation*. Leuven: Peeters.

Haig, Geoffrey L. J. 2002. Complex predicates in Kurdish: Argument sharing, incorporation, or what? *STUF – Language Typology and Universals* 55(1). 15–48.

Haig, Geoffrey L. J. 2007. Grammatical borrowing in Kurdish (Northern group). In Yaron Matras & Janet Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 165–184. Berlin: Mouton de Gruyter.

Haig, Geoffrey L. J. 2011. Linker, relativizer, nominalizer, tense-particle: On the *ezafe* in West Iranian. In Foong Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages: Diachronic and typological perspectives*, vol. 1: *Sino-Tibetan and Iranian languages*, 363–390. Amsterdam: John Benjamins.

Haig, Geoffrey L. J. & Ergin Öpengin. 2014. Kurdish: A critical research overview. *Kurdish Studies* 2(2). 99–122.

Haig, Geoffrey L. J. & Ergin Öpengin. 2015. Gender in Kurdish: Structural and socio-cultural dimensions. In Marlis Hellinger & Heiko Motschenbacher (eds.), *Gender across languages*, vol. 4, 247–276. Amsterdam: John Benjamins.

Haig, Geoffrey L. J. & 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, Geofrey. 2014. East Anatolia as a linguistic area? Conceptual and empirical issues. In Lale Behzadi, Patrick Franke, Geofrey Haig, Christoph Herzog, Birgitt Hofmann, Lorenz Korn & Susanne Talabardon (eds.), *Bamberger Orientstudien*, 13–36. Bamberg: University of Bamberg Press.

Ibrahim, Iman. 2015. Gender assignment to lexical borrowings by heritage speakers of Arabic. *Western Papers in Linguistics / Cahiers linguistiques de Western* 1(2).

James, Boris. 2007. Le «territoire tribal des Kurdes» et l'aire iraqienne (Xe-XIIIe siècles): Esquisse des recompositions spatiales. *Revue des mondes musulmans et de la Méditerranée* 10(117–118). 101–126.

Jastrow, Otto. 2011. Turkish and Kurdish influences in the Arabic dialects of Anatolia. *Türk Dilleri Araştırmaları* 21(1). 83–94.

Kahn, Margaret. 1976. *Borrowing and regional variation in a phonological description of Kurdish*. Ann Arbor, MI: Phonetics Laboratory of the University of Michigan.

Kent, Roland. 1953. *Old Persian*. 2nd, revised edn. New Haven, CT: American Oriental Society.

Lucas, Christopher. 2015. Contact-induced language change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.

MacKenzie, David N. 1961. *Kurdish dialect studies*. Vol. 1. Oxford: Oxford University Press.

Matras, Yaron. 2010. Contact, convergence, and typology. In Raymond Hickey (ed.), *The handbook of language contact*, 66–85. Oxford: Wiley-Blackwell.

Matras, Yaron & Andrew Koontz-Garboden. 2016. *The dialects of Kurdish*. University of Manchester. http://kurdish.humanities.manchester.ac.uk/.

Matras, Yaron & Jeanette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31. 829–865.

Noorlander, Paul M. 2014. Diversity in convergence: Kurdish and Aramaic variation entangled. *Kurdish Studies* 2(2). 201–224.

Öpengin, Ergin. 2012. Sociolinguistic situation of Kurdish in Turkey: Sociopolitical factors and language use patterns. *International Journal of the Sociology of Language* 2012(217). 151–180.

Öpengin, Ergin. 2016. *The Mukri variety of Central Kurdish: Grammar, texts, and lexicon*. Wiesbaden: Reichert.

Öpengin, Ergin. Forthcoming. The history of Kurdish and the development of literary Kurmanji. In Hamit Bozarslan, Cengiz Gunes & Veli Yadirgi (eds.), *Cambridge history of the Kurds*. Cambridge: Cambridge University Press.

Öpengin, Ergin & Geoffrey L. J. Haig. 2014. Regional variation in Kurmanji: A preliminary classification of dialects. *Kurdish Studies* 2(2). 143–176.

Paul, Ludwig. 2008. Kurdish language i: History of the Kurdish language. In Ehsan Yarshater, Mohsen Ashtiany & Mahnaz Moazami (eds.), *Encyclopædia iranica*, online edn. New York: Encyclopædia Iranica Foundation.

Perry, John R. 2005. Lexical areas and semantic fields of Arabic loanwords in Persian and beyond. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal divergence: Case studies from Iranian, Semitic and Turkic*, 97–110. London: Routledge.

Perry, John R. 2012. New Persian: Expansion, standardization, and inclusivity. In Brian Spooner & William L. Hanaway (eds.), *Literacy in the Persianate world:*

Writing and the social order, 70–94. Philadelphia: University of Pennsylvania Museum of Archaeology & Anthropology.

Ripper, Thomas. 2012. *Diyarbekir merwanileri: İslami ortaçağ'da bir Kürt hanedanı*. Istanbul: Avesta. Translated by Bahar Şahin Fırat. Originally published in 2000 by Ergon.

Ryding, Karin C. 2014. *Arabic: A linguistic introduction*. Cambridge: Cambridge University Press.

Sirkeci, Ibrahim. 2005. War in Iraq: Environment of insecurity and international migration. *International Migration* 43(4). 197–214.

Teyran, Feqiyê. 2011. *Dîwana Feqiyê Teyran*. Seîd Dêreşî (ed.). Diyarbakir: Lîs.

Todd, Terry L. 2002. *A grammar of Dimili (also known as Zaza)*. 2nd edn. Stockholm: Iremet.

Tsabolov, Ruslan. 1994. Notes on the influence of Arabic on Kurdish. *Acta Kurdica* 1. 121–124.

Utas, Bo. 2005. Semitic in Iranian: Written, read and spoken language. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal divergence: Case studies from Iranian, Semitic and Turkic*, 65–77. London: Routledge.

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.

Zeyneloğlu, Sinan, Ibrahim Sirkeci & Yaprak Civelek. 2016. Language shift among Kurds in Turkey: A spatial and demographic analysis. *Kurdish Studies* 4(1). 25–50.

Chapter 9

Jerusalem Domari

Yaron Matras
University of Manchester

Jerusalem Domari is the only variety of Domari for which there is comprehensive documentation. The language shows massive influence of Arabic in different areas of structure – quite possibly the most extensive structural impact of Arabic on any other language documented to date. Arabic influence on Jerusalem Domari raises theoretical questions around key concepts of contact-induced change as well as the relations between systems of grammar and the components of multilingual repertoires; these are dealt with briefly in the chapter, along with the notions of fusion, compartmentalisation of paradigms, and bilingual suppletion.

1 Historical development and current state

Domari is a dispersed, non-territorial minority language of Indo-Aryan origin that is spoken by traditionally itinerant (peripatetic) populations throughout the Middle East. Fragmented attestations of the language place it as far north as Azerbaijan and as far south as Sudan. The self-appellation *dōm* is cognate with those of the *řom* (Roma or Romanies) of Europe and the *lom* of the Caucasus and eastern Anatolia. All three populations show linguistic resources of Indo-Aryan origin (which in the case of the Lom are limited to vocabulary), as well as traditions of a mobile service economy, and are therefore all believed to have descended from itinerant service castes in India known as *ḍom*. Some Domari-speaking populations are reported to use additional names, including *qurbāṭi* (Syria and Lebanon), *mıtrıp* or *karači* (Turkey and northern Iraq) and *bahlawān* (Sudan), while the surrounding Arabic-speaking populations usually refer to them as *nawar*, *ɣağar* or *miṭribiyya*. The language retains basic vocabulary of Indo-Aryan origin, and shows elements of lexical phonology that place its early development within the Central Indo-Aryan group of languages. It retains conservative derivational as

Yaron Matras. 2020. Jerusalem Domari. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 203–227. Berlin: Language Science Press. DOI:??

well as present-tense inflectional verb morphology that goes back to late Middle Indo-Aryan, alongside innovations in nominal and past-tense verb inflection that suggest that the language was contiguous with the Northwestern frontier languages (Dardic) during the transition to early modern Indo-Aryan (cf. Matras 2012).

The first attestation of Palestinian Domari is a list of words and phrases collected by Ulrich Jasper Seetzen in 1806 in the West Bank and published by Kruse (1854). It was followed by Macalister's (1914) grammatical sketch, texts and lexicon, collected in Jerusalem in a community which at the time was still nomadic, moving between the principal West Bank cities of Nablus, Jerusalem and Hebron. This community settled in Jerusalem in the early 1920s, the men taking up wage employment with the British-run municipal services. In the 1940s they abandoned their makeshift tent encampment and moved into rented accommodation within the Old City walls, where the community still resides today. Between 1996 and 2000 I carried out fieldwork among speakers in Jerusalem and published a series of works on the language, including two descriptive outlines (Matras 1999; 2011), annotated stories (Matras 2000), an overview of contact influences (Matras 2007), and a descriptive monograph (Matras 2012).

A number of sources going back to Pott (1907), Newbold (1856), Paspati (1870), Patkanoff (1907), and Black (1913) provide language samples collected among the Dom of Lebanon, Syria, Iraq and the Caucasus. These are supplemented by a few more samples collected by ethnographers (cf. Matras 2012: 15ff.) and subsequently by data collected in Syria and Lebanon by Herin (2012). That documentation allowed me to identify a number of differences that appeared to separate a Northern group of Domari dialects from a Southern group, which latter includes the data recorded in Palestine as well as a sample from Jordan (see Matras 2012: 15ff.). That tentative classification has since been embraced by Herin (2014), who goes a step further and speculates about an early split between two branches of the language. To date, however, published attestation of Northern varieties remains extremely fragmented, notwithstanding recent work by Herin (2016; this volume), while the only comprehensive overview of a Southern variety remains that from Jerusalem.

Outside of Jerusalem and its outskirts there are known communities of Palestinian Doms in some of the refugee camps on the West Bank and Gaza, as well as in Amman, where a few families sought refuge in 1967. Numbers of speakers were very low in all these communities already in the mid 1990s and the language was only in use among the elderly. During my most recent visit to the Jerusalem community, in January 2017, it appeared that there was only one sin-

gle fluent speaker left, who, for obvious reasons, no longer had any practical use for the language, apart from flagging the odd phrase to younger-generation semi-speakers. Jerusalem Domari, and most likely Palestinian Domari in general, must therefore now be considered to be nearly extinct.

2 Contact languages

Given the migration route that the Dom will have taken to reach the Middle East from South Asia, it is plausible that the language was subjected to repeated and extensive contact influences. Kurdish influences on Jerusalem Domari, some of them attributable specifically to Sorani Kurdish, and some Persian items, are apparent in vocabulary, while some of the morpho-syntactic structures (such as extensive use of person affixes, and the use of a uniform synthetic marker of remote tense that is external to the person marker) align themselves with various Iranian languages. There is also a layer of Turkic loans, some of which may be attributable to Azeri varieties, while others are traceable to Ottoman rule in Palestine; such items are numerous in the wordlists compiled by Seetzen and Macalister during the Ottoman period, but are much less frequent in the materials collected a century later (for a discussion of etymological sources see Matras 2012: 426–429).

The circumstances under which speakers of Domari first came into contact with Arabic are unknown. There are some indications of a layered influence: Domari tends to retain historical /q/ in Arabic-derived words, as in *qahwa* 'coffee', *qabil* 'before', *qaddēš* 'how much', as found in the rural dialects of the West Bank (and elsewhere), whereas contemporary Jerusalem Arabic (also used by Doms when speaking Arabic) shows a glottal stop, as in *ʔahwe*, *ʔabl, ʔaddēš*; the word for 'now' is *hessaʕ*, while Jerusalem Arabic has *hallaʔ*. It appears that the community has been fully bilingual in Arabic and Domari at least since the early 1800s, with knowledge of Turkish having been widespread among adults during the Ottoman rule. Due to the nature of the Doms' service economy, Arabic was an essential vehicle of all professional life, whether metalwork, hawking, begging, or performance, but Domari remained the language of the household until the introduction of compulsory school education under Jordanian rule in the 1950s–60s, at which point parents ceased to pass on the language to children. By the 1990s, use of Domari was limited to a small circle of perhaps around forty–fifty elderly people. Due to the multi-generational structure of households it was rare even then for conversations to be held exclusively among Domari speakers. Domari–Arabic bilingualism has always been unidirectional, with Arabic being

the language of commerce and public interactions for all Doms, and more recently also of education and media, eventually replacing Domari as a home and community language.

3 Contact-induced changes in Jerusalem Domari

As a result of ubiquitous bilingualism among all Domari speakers, Domari talk is chequered not only with expressions that derive from Arabic, but also with switches into Arabic for stylistic and discourse-strategic purposes such as emphasis, direct quotes, side remarks, and so on. The structural intertwining of Domari and Arabic, and the degree to which active bilingual speakers maintain a license to incorporate Arabic elements into Domari conversation, pose a potential challenge to the descriptive agenda. In the following I discuss those structures that derive from Arabic, and are shared with Arabic (in the sense that they are employed by speakers both in the context of Domari conversation and in interactions in Arabic) but constitute a stable and integral part of the structural inventory of Domari without which Domari talk cannot be formed, and for which there is no non-Arabic Domari alternative. All examples are taken from the Jerusalem Domari corpus described in Matras (2012). Examples from Arabic are based on colloquial Palestinian Arabic as spoken in Jerusalem.

3.1 Phonology

The entire inventory of Palestinian Arabic phonemes is available in Domari; Arabic-derived words that are used in Domari conversation (whether or not they have non-Arabic substitutes) do not undergo phonological or phonetic integration, except for the application of Domari grammatical word stress on case-inflected nouns (e.g. *lambá* 'lamp.ACC', from Arabic *lámba*). The pharyngeals [ḥ] and [ʕ] are limited to Arabic-derived vocabulary. The sounds [q], [ɣ] and [ɬ] as well as [z] and [f] appear primarily in Arabic-derived vocabulary, but there is evidence that they entered the language already through contact with Turkic and Iranian languages. Less clear is the status of the pharyngealised dental consonants /ḍ, ṭ, ṣ/. These are largely confined to Arabic-derived vocabulary, but they can also be found in inherited words of Indo-Aryan stock, where they often represent original (Indo-Aryan) retroflex sounds (cf. *ḍōm* 'Dom', *pēṭ* 'belly'). An ongoing phonological innovation that is shared with Jerusalem Arabic is the simplification of the affricate [ʤ] to the fricative [ʒ] in inherited lexemes, e.g. *džami* 'I go' > *žami*. This triggers a corresponding simplification of [ʧ] to [ʃ], as in *lači* 'girl' > *laši*.

3.2 Morphology

Domari has not adopted productive word-derivational templates from Arabic. Arabic inflectional morphology, however, is productive with some Arabic-derived word forms, resulting, in effect, in a compartmentalised morphological structure. Arabic-derived plural nouns tend to retain Arabic plural inflection, but indigenous (inherited, Indo-Aryan) plural inflections are added to the word: thus *muslim* 'Muslim', plural *musilmīn-e* Muslims-PL 'muslims'; *madrase* 'school', dative plural *madāris-an-ka* (schools-PL.OBL-DAT) 'to the schools'. While Jerusalem Domari retains inherited plural marking with nouns derived from both Indo-Aryan and Arabic, in the closely related variety of the nomadic Doms of Jordan the Arabic plural ending *-āt* is often used with inherited nouns: thus *putur* 'son', Jerusalem Domari plural *putr-e*, Jordanian Domari plural *putr-āt*.

Arabic person agreement inflection is retained with Arabic-derived modal and aspectual auxiliaries. The auxiliaries *kān* 'be', *ṣār* 'begin', and *baqa* 'continue' take Arabic verbal inflection, while *bidd-* 'want', *ḍall-* 'continue', and *ḫallī-* 'allow' take Arabic nominal-possessive marking:

(1) a. kān-at par-ar-m-a wāšī-s
be.PRF-3SG.F take-3SG-1SG-PST with-3SG
'She used to take me with her.'

b. dōm-e kān-u kam-k-ad-a ḥaddādīn-e
dom-PL be-.PRF-3PL work-TR-3PL-PST blacksmiths-PL
'The Dom used to work as blacksmiths.'

(2) a. ṣār qaft-ar-i min ɔy-os
begin.PRF.3SG.M steal-3SG-PROG from father-3SG
'He started to steal from his father.'

b. ṣār-u kar-and-i ḥafl-e
begin.PRF-3PL do-3PL-PROG party-PL
'They started to have parties.'

(3) a. š-ird-i ama-ke bidd-ha qumn-ar
say-PFV-F 1SG-BEN want-3SG.F eat-SBJV.3SG
'She said to me that she wants to eat.'

b. bidd-i par-am itžawwiz-om-is
want-1SG take-1SG.SBJV marry-1SG.SBJV-3SG.OBL
'I want to take her and marry her.'

(4) a. ḫallī-hum naḍḍif-k-ad-i ehe marn-an
let.IMP.2SG-3PL clean-TR-3PL-PROG these.PL dead-OBL.PL
'Let them clean up these corpses.'

b. ḫallī-h rʕi-k-ar hundar
let.IMP.2SG-3SG graze-TR-3SG.SBJV there
'Let it graze there.'

Inflected Arabic-derived auxiliaries include the existential verb *kān-* 'to be', which is used in Domari, as in Arabic, as a past- and future-tense copula, supplementing the Domari remoteness or 'external' past-tense marker *-(y)a*, which follows the lexical predication or predicate object:

(5) ihi illi par-d-om-is kān-at yatīm-ēy-a
this.F REL take-PST-1SG-3SG.OBL be.PRF-3SG.F orphan-PRED.SG-PST
'The one [woman] whom I married [her] was an orphan.'

Arabic-derived auxiliaries are also inflected for tense following Arabic paradigms:

(6) lāzem tkūn itme mišaṭṭaṭ-hr-es-i
must be.IMPF.SBJV.3SG.F 2PL dispersed-ITR-2PL-PROG
'You must remain dispersed.'

This amounts, in effect, to a functional compartmentalisation in verbal morphology: both inherited and Arabic-derived lexical verbs take inherited Indo-Aryan inflection, while Arabic-derived modal and aspectual auxiliaries take Arabic inflection (for further discussion see Matras 2015).

Arabic person inflection is also found with the Arabic-derived secondary pronominal object marker *iyyā-*, complementiser *inn-*, and conjunction *liʔann-* 'because':

(7) ple illi t-or-im iyyā-hum
money.PL REL give.PST-2SG-1SG.OBL OBJ-3PL
'the money that you gave [it] to me'

(8) aɣlabiyy-osan š-ad-i inn-hom min šamāl-os-ki
majority-3PL say-3PL-PROG COMP-3PL from north-3SG-ABL
hnūd-an-ki
india-OBL.PL-ABL
'Most of them say that they are from northern India.'

(9) na kil-d-om barra liʔann-ha wars-ar-i
NEG exit-PFV-1SG out because-3SG.F rain-3SG-PRS
'I did not go out because it was raining.'

(10) payy-os liʔinn-o ṭāṭ-i kān
husband-3SG because-3SG.M Arab-PRED.SG be.PRF.3SG.M
'Because her husband was an Arab.'

Note that in example (9) the agreement is in the feminine singular, corresponding to the grammatical mapping of the Jerusalem Arabic construction 'it rains' where the (underlying) subject is the feminine noun *dunya* 'the world', while in (7), resumptive pronoun agreement with 'money', a plural noun, is in the plural.

Domari is seemingly an exception to the frequently cited generalisation that derivational morphology is more likely to be borrowed than inflectional morphology (cf. Moravcsik 1978; Field 2002; Matras 2009: §6.2.2). In fact, the constraint on the borrowing of word-derivational morphology results from the clash with the principle of the transparency of morphemes (cf. Matras 2009: §6.2.2): Arabic has few if any word-derivational morphemes that can be isolated, relying instead on complex morphological templates into which lexical roots are inserted. Nominal plural morphemes have both inflectional function (relevant to other elements in the clause) and derivational function (having independent meaning in standalone expressions). As shown above, they are replicated in Jerusalem Domari as an integral part of Arabic plural word forms. On the other hand, the replication of inflectional material on auxiliaries is not productive, in that it is not incorporated into the general lexicon, not even with lexical words of Arabic origin, but remains confined to the near-wholesale adoption of modal and aspectual auxiliaries from Arabic. In this respect, Arabic-derived inflectional paradigms in Domari constitute a case of both FUSION as defined in Matras (2009) – the wholesale non-separation of language systems around a particular functional category – and at the same time a case of functional compartmentalisaton as defined in Matras (2015) – the distinct treatment of functional sub-components of a category, here the verbal category, in regard to grammatical inflection.

3.3 Syntax

Generally, Jerusalem Domari shows full congruence with Palestinian Arabic in most syntactic functions. This includes word order rules and the formation of both simple and complex clauses. It also includes configurations such as mapping

of tenses and modality to complement and conditional clauses, and the mapping of semantic relations onto case markers. The latter can be adpositional or inflectional. For nominal possessive constructions, Domari has two options. The first of those options, illustrated in (11a), is what we might call canonical Domari. It corresponds to the inherited Indo-Aryan pattern. The second option, illustrated in (11b), corresponds to the common Palestinian Arabic construction, which is presented in (11c). Here Domari replicates the role of the Arabic dative preposition *la* by means of the inherited Domari ablative/possessive inflectional ending *-ki*:

(11) a. Canonical Domari
bɔy-im kuri
father-1SG house

b. Convergent Domari
kury-os bɔy-im-ki
house-3SG father-1SG.OBL-ABL

c. Arabic
bēt-o la-ʔabū-y
house-3SG.M to-father-OBL.1SG
'my father's house'

The canonical position of adjectives in Domari is, as in other Indo-Aryan languages, before the noun (12a), while in Arabic adjectives follow the noun. However, speakers show an overwhelming preference for avoiding pre-posed adjectives and instead make use of the non-verbal predication marker in order to allow the adjective to follow the noun (12b), thereby replicating Arabic word order patterns (12c):

(12) a. Canonical Domari
er-i qišṭoṭ-i šōni
come.PFV-F little-F girl
'A little girl arrived.'

b. Convergent Domari
er-i šōni qišṭoṭ-ik
come.PFV-F girl little-PRED.SG.F
'A little girl arrived.' [= 'A girl arrived, being little.']

c. Arabic
ʔižat bint zɣīre
come.PRF.3SG.F girl little.F
'A little girl arrived.'

The emergence of nominal clauses, facilitated by the availability of non-verbal predication markers, might be regarded as an innovation for an Indo-Iranian language, which reinforces sentence-level convergence between Arabic and Domari:

(13) a. Domari
wuda bizzot-ēk
old.M poor-PRED.SG

b. Arabic
l-ḫityār miskīn
DEF-old.man poor
'The old man is poor.'

Domari, like Arabic, shows a strong tendency toward SVO word order in categorical sentences in which a thematic perspective is established by linking to a known topical entity:

(14) mām-om putur yāsir gar-a swēq-ē-ta
uncle-1SG son Yassir go.PFV-M market-OBL.F-DAT
'My (paternal) cousin Yassir went to the market.'

By contrast, as seen in example (12), Domari shows consistent convergence with Arabic in regard to the position of the subject after the verb when new topical entities are introduced, especially with verbs that convey movement and change of state and in presentative constructions. Drawing on inherited morphology, this convergence in word order patterns also allows for the encoding of the pronominal experiencer–recipient through a person affix that is attached to an intransitive verb in presentative constructions, matching the Arabic construction:

(15) a. Domari
er-os-im ḫabar
come.PFV-3SG-1SG.OBL notice
b. Arabic
ʔažā-ni ḫabar
come.PRF.3SG.M-1SG notice
'I received notification'

Complex clauses are also congruent with Arabic. Like Arabic, Domari shows three distinct co-temporal adverbial constructions. In the first, the subordinate clause is introduced by the conjunction 'and' and the verb is finite and indicative:

(16) a. Domari
kahind-ad-i ū pandži našy-ar-i
look-3PL-PROG and 3SG dance-3SG-PROG
b. Arabic
b-yitfarražu w hiyye b-turʔuṣ
IND-look.IMPF.3PL and 3SG.F IND-dance.IMPF.3SG.F
'They watch her dance.'

In the second, the subordinated predicate appears in the present participle:

(17) a. Domari
lah-erd-om-is mindir-d-ēk
see-PFV-1SG-3SG.OBL stand-PFV-PRED.SG.M
b. Arabic
šuft-o wāʔef
see.PRF.1SG-3SG.M standing
'I saw him standing.'

The final option shows a nominalised verb, whose possessive inflection indicates the subject/agent, introduced by the preposition 'with' in the subordinate position alongside a finite main clause:

(18) a. Domari
maʕ šuš-im-ki tiknaw-ar-m-i gurg-om
with sleep-1SG.OBL-ABL hurt-3SG-1SG-PROG neck-1SG

b. Arabic
maʕ nōmt-i b-tūžaʕ-ni raʔbt-i
with sleep-OBL.1SG IND-3SG.F-hurt.IMPF.3SG.F-1SG neck-OBL.1SG
'As I sleep, my neck hurts.'

Relative clauses follow the format of Arabic relative clauses: they employ the Arabic-derived post-nominal relativiser *illi* and show the same distribution rules for pronominal resumption as in Arabic:

(19) ihi illi par-d-om-is kān-at yatīm-ēy-a
this.F REL take-PST-1SG-3SG.OBL be.PRF-3SG.F orphan-PRED.SG-PST
'The one [woman] whom I married [her] was an orphan.'

Factual (indicative) complements are introduced by the Arabic-derived complementiser *inn-*, which carries Arabic-derived inflection (as in example 8 above), and show comparable clause structure as in Arabic:

(20) a. Domari
džan-ad-i in-na dōm
know-3PL-PROG COMP-1PL Dom
b. Arabic
b-yiʕrafu in-na dōm
IND-know.IMPF.3PL COMP-1PL Dom
'They know that we are Dom.'

Modal complements and same-subject purpose clauses show, as in Arabic, a subjunctive complement, without a complementiser:

(21) a. Domari
bidd-i dža-m ḥaram-ka ṣalli-k-am
want-1SG go-1SG.SBJV mosque-DAT pray-TR-1SG.SBJV
b. Arabic
bidd-i arūḥ ʕa-l-ḥaram aṣalli
want-1SG go.IMPF.SBJV.1SG to-DEF-mosque pray.IMPF.SBJV.1SG
'I want to go to the mosque to pray.'

Adverbial clauses employ Arabic-derived adverbial subordinators, including *lamma* 'when', as in (22), or composite conjunctions consisting of a preposition and complementiser, such as *baʕd mā* 'after' and *qabil mā* 'before', as in (23) and (24), and generally follow Arabic sentence organisation and tense and modality distribution patterns.

(22) lamma lak-ed-a ḫāl-os inǧann-ahr-a bɔy-om
when see-PFV-M uncle-3SG crazy-TR.PFV-M father-1SG
'When he saw his uncle, my father went crazy.'

(23) baʕd mā ḫaḷḷaṣ-k-ed-a kam-os gar-a kury-is-ta
after COMP finish-TR-PFV-M work-3SG go.PFV-M house-3SG.OBL-DAT
'After he finished his work he went home.'

(24) qabil mā dža-m ḫaḷḷaṣ-k-ed-om kam-as
before COMP go-1SG.SBJV finish-TR-PFV-1SG work-OBL.M
'Before I left I finished my work.'

Conditional clauses similarly draw on the Arabic conjunctions *iza* and *law*, both 'if', and show similar distribution of tense and aspect categories, including the Arabic-derived impersonal marker of counter-factuality *kān,* literally 'was':

(25) a. Domari
law er-om ḫužoti kān lah-erd-om-s-a
if come.PFV-1SG yesterday was see-PFV-1SG-3SG-PST
b. Arabic
law žīt mbāreḥ kān šuft-o
if come.PRF.1SG yesterday be.3SG.M see.PRF.1SG-3SG.M
'If I had come yesterday, I would have seen him.'

3.4 Lexicon

Jerusalem Domari shows extensive impact of Arabic on the grammatical lexicon, including almost wholesale reliance on Arabic-derived material for entire categories. In the pronominal domain, Domari employs, in additional to the secondary pronominal object marker *iyyā-* discussed above, also the Arabic reflexive pronoun *ḥāl-*, derived from the word 'state', combined with person/possessive inflection, and the Arabic reciprocal pronoun *baʕḍ-*:

(26) naḍḍif-k-ad-a ḥāl-os
clean-TR-PFV-M REFL-3SG
'He cleaned himself.'

(27) tʕarraf-h-r-ēn baʕḍ-ē-man-ta
meet-TR.PFV-1PL RECP-PL-1PL-DAT
'We met one another.'

Indefinite expressions draw on Arabic-derived forms of category determination including negative *wala*, free choice *ayy* and universal *kull*, which may be combined with inherited ontological markers, as well as on the ontological specifiers *ḥāğ-* for thing and *maḥall* for location. Indefinite expressions that derive entirely from Arabic include temporal *wala marra* 'never', *dāyman* 'always', and universal-thing *kullši* 'everything'. Arabic-derived focus particles are *barḍo* 'also, too' and *ḥatta* 'even' and quantifiers are *kull* 'every, each' and *akamm* 'a few'. Interrogatives are generally inherited (Indo-Aryan), with the exception of *qaddēš* 'how much'. Numerals are all derived from Arabic with the exception of the lowest numeral forms ('one' to 'five' in citation function and 'one' to 'three' in attributive role) (see Tables 1–2); all ordinal numerals (*awwal* 'first', *tāni* 'second' etc.) are from Arabic.

Alongside a very small number of inherited prepositions that are used exclusively with pronominal (person-inflected) forms, most prepositions are derived from Arabic (Table 3).

Arabic-derived grammatical operators at verbal clause level include a series of modality adverbs such as *masalan* 'for example', *yimken* 'perhaps', *atāri* 'well', time adverbs such as *hessaʕ* 'now' and *baʕdēn* 'then, afterwards', and the phasal adverbs *lissa* and *lāyzāl*, both 'still'. As discussed above, Domari adopts Arabic modal and aspectual auxiliaries wholesale, i.e. along with their Arabic-derived inflection. This covers almost the full category of modal and aspectual auxiliaries including habitual/iterative *kān* 'be', *ṣār* 'begin', and *baqa* 'continue', *bidd-* 'want', *ḍall-* 'continue', and *ḫallī-* 'let', as well as the impersonal form *lāzem* 'must'. The only modal for which an Indo-Aryan form is retained is *sak-* 'to be able to'. Past-tense finite predications take the Arabic negator *mā* (alongside inherited *na*) while in non-finite predications the Arabic negation particle *miš* is used:

(28) mā lak-ed-om-is
NEG see-PFV-1SG-3SG.OBL
'I didn't see him/her.'

(29) bay-os miš kury-a-m-ēk
wife-3SG NEG house-OBL.F-LOC-PRED.SG
'His wife is not at home.'

Table 1: Jerusalem Domari numerals

Numeral	Citation	Attribute
1	*ikak*	*-ak*
2	*diyyes*	*di*
3	*taranes*	*taran*
4	*štares*	*ʔarbaʃ*
5	*pʌndžes*	*ḫamis*
6	*sitt-ēk-i*	*sitt*
7	*sabʃ-ak-i*	*sabaʃ*
8	*tamāni-ak-i*	*tamānye*
9	*tisʃ-ak-i*	*tisʃa*
10	*das* ‘ten’, *ʃašr-ak-i*	*ʃašr*
20	*ʃišrīn-i, wīs-i*	*ʃišrīn*
21	*ʃišrīn ū ekak-i*	*wāḥed w ʃišrīn*
22	*ʃišrīn-i ū diyyes-i*	*tnēn w ʃišrīn*
23	*ʃišrīn-i ū taranes-i*	*talāte w ʃišrīn*
24	*ʔarbaʃ ū ʃišrīn*	*ʔarbaʃ w ʃišrīn*
100	*miyyēk hi, siyy-ak-i*	*miyye*
1000	*alf-ak-i*	*alf*

Table 2: Jerusalem Domari higher numerals

Numeral	Form
30	*talātīn*
40	*ʔarbaʃīn*
50	*ḫamsīn*
60	*sittīn*
70	*sabʃīn*
80	*tamanīn*
90	*tisʃīn*

Table 3: Arabic-derived prepositions in Jerusalem Domari

ʕan	'on, about'	*ʕašān*	'because'	*nawāḥi*	'toward'
maʕ	'with'	*minšān*	'for'	*qabil*	'before'
min	'from'	*min ɣēr*	'without'	*baʕd*	'after'
la, ʕala	'to'	*min/bi dūn*	'without'	*layāyet*	'until'
fi	'in'	*bēn*	'between'	*bi*	'in, for'
zayy	'like'	*ḥawāli*	'around'	*ḍiḍḍ*	'against'
ʕind (ʕand)	'at'	*min ḍamn*	'among'	*žamb*	'next to'
badāl	'instead of'	*ʔilla ɣēr*	'except for'		

Clause combining relies exclusively on Arabic-derived material (connectors and conjunctions) (see Table 4).

Table 4: Arabic-derived conjunctions in Jerusalem Domari

w	'and'	*qabil mā*	'before'
wala	'and not', '(n)either'	*baʕd mā*	'after'
yā	'or'	*min-yōm-mā*	'since'
willa	'or (else)'	*iza*	'if'
bass	'but', 'only'	*law*	'if'
illi	relative pronoun	*bi-r-rayim*	'despite'
inn-	'that'	*ʕašān*	'for', 'in order to'
liʔann	'because'	*minšān*	'for', 'in order to'
lamma	'when'	*ta*	'in order to'
kull mā	'whenever'		

Likewise, the inventory of discourse particles and interjections is adopted in its entirety from Arabic: We find the interjection, tags and filers *yabayyi, yaḷḷa, xaḷaṣ, waḷḷa,* and *yaʕni*, as well as segmental markers with a lexical meaning such as *l-muhimm* 'anyway', *l-ḥāṣil* 'finally', *ṭayyib* 'well', *w ʔiši* 'and the like', *w hāda* 'and so on', *abṣar* 'whatever', and the filler *hāy* 'that'. The quotation particle *qal/ḥal*, from Arabic 'say', is not found in Jerusalem Arabic and appears to represent an older layer of Arabic influence (as indicated also by its phonological structure; see §2).

The content lexicon equally shows massive impact of Arabic. In the Jerusalem Domari corpus of narrational and conversational talk as well as sentence elicitation recorded in the 1990s (Matras 2012), almost two thirds of lexical items are Arabic-derived; the count includes single-word insertions from Arabic, including attributive nominal compounds (noun–possessor and noun–adjective), but excludes phrases containing a finite lexical verb that is Arabic-derived (which latter are regarded as optional code-switches). Both Arabic-derived nouns and adverbs outnumber inherited (Indo-Aryan) counterparts by around 65% to 35%, while for verbs and adjectives the numbers are roughly equal. Around 26% of items of both the Swadesh 100-item list and the Leipzig–Jakarta 100-item list (Haspelmath & Tadmor 2009) are Arabic-derived. This puts Domari in the range of languages considered to be "high borrowers" by the Leipzig Loanword Typology Project (Haspelmath & Tadmor 2009). Meanings on the list that are replaced by Arabic loans in Domari include a number of animals ('ant', 'bird', 'fish'), activities ('to run', 'to fly', 'to crush'), elements of nature ('star', 'soil', 'shade', 'ash', 'leaf', 'root'), and some body parts ('knee', 'navel', 'liver', 'thigh'; also 'wing', 'tail'). On the whole, the meaning and usage of Arabic-derived lexemes matches that of Jerusalem Arabic. Creative processes are marginal and include such processes as the phonological volatility of /q/ (as [q], [x], [qx] and [g]), the alternation between *farğik-* 'to show' (Arabic *√frğ*) and *warğik-*, and the occasional creative derivation such as *bisawahr-* 'to get married', from Arabic *bi-sawa* 'together'.

Arabic verbs are integrated into Domari through a light verb construction that draws on the inherited verb stems *-k-* 'to do' and *-h-* 'to become', which are grammaticalised into loan-verb adaptation markers (see Matras 2012: 240–244) that are sensitive to valency. This follows a strategy for the adaptation of loan verbs that is widespread across a geographical area stretching from the Balkans and the Caucasus through Anatolia and Western Asia and on to the Indian Subcontinent. For some verbs, alternating adaptation markers can indicate change in valency: *ğawwiz-h-r-i* (marry-ITR-PFV-F) 'she got married', *ğawwiz-k-am-is* (marry-TR-1SG.SBJV-3SG.OBL) 'I shall marry her off'. The core of integrated Arabic verbs generally derives from the Arabic subjunctive–imperative form, which in Arabic never occurs in isolation from its person inflection in the prefix conjugation, as in *ğawwiz-* 'marry', from *ğawwiz 'marry (off)!' or *tğawwiz 'get married!'. Note, however, that the vowel structure of the core does not always correspond to the subjunctive–imperative form of contemporary Palestinian Arabic, which is quite possibly a further indication of the layered historical influence of Arabic. Thus we find *s'il-k-ed-om* (ask-TR-PFV-1SG) 'I asked', from *s'il- 'ask', while Palestinian Arabic has *is?al* 'ask!', and *rawwaḥ-ah-r-a* (go-ITR-PFV-M) 'he travelled', while Palestinian Arabic has *rawwiḥ* 'go away!'.

3.5 Cross-category interplay

A typologically curious case of contact-induced change is offered by the use in Jerusalem Domari of three construction types that cut across structural categories. The first pertains to the comparative form of adjectives. In the absence of a structurally transparent, isolated and replicable marker of adjective comparison (comparative and superlative), Domari draws on Arabic word forms for all comparative adjective forms, even when an inherited (non-Arabic) word form is used for the positive form of the adjective, as illustrated in (30) (cf. Herin, this volume: §3.2).

(30) a. Domari
atu qaštot-ik
you.SG small-PRED.SG.F
'You are small.'

b. Domari
atu azɣar mēšī-m-i
you.SG smaller from-1SG-PRED.SG
'You are smaller than I.'

c. Arabic
inti zɣīre
2SG.F small.F
'You are small.'

d. Arabic
inti azɣar minn-i
2SG.F smaller from-OBL.1SG
'You are smaller than I.'

This formation involves essentially the recruitment of an alternative, Arabic-derived item from the category of lexical items in order to carry out a grammatical procedure that is derivational–inflectional by nature (derivational in that it modifies meaning, inflectional in that it is inherently embedded into a syntactic relationship at the phrase level); thus we have a case of cross-category interplay.

A further case is that of lexical suppletion around Arabic-derived numerals. Domari and Arabic differ typologically in respect of numeral agreement: with Indo-Aryan numerals, the Domari noun appears in the default singular form, while in Arabic, numerals up to 'ten' take plural agreement. The clash is resolved in Domari in such a way that Arabic-derived numerals under 'ten' invariably

trigger an Arabic-derived lexical item even when an inherited form of the corresponding lexeme is available:

(31) a. ḫkum-ke-d-os taran wars maḥkame
sentence-TR-PFV-3SG three year court
'The court sentenced him to three years.'

b. eh-r-a ʕumr-om sitte snīn
become-PFV-M age-1SG six year.PL
'I turned six years old.'

Such alternation is systematic (see further examples in Table 5) and might be regarded as a case of bilingual suppletion, where every countable noun in the language for which an inherited (Indo-Aryan) word form exists also has an Arabic-derived counterpart that is used with numerals between 'three' and 'ten'.

Table 5: Some phrases from the corpus containing numerals and nouns

Inherited numeral and singular noun	Arabic numeral and plural noun
di dīs taran dīs 'two days three days'	*sabaʕ-t-iyyām* 'seven days'
taran mas 'three months'	*ḫamas-t-ušhur* 'five months'
taran wars 'three years'	*sitte snīn* 'six years'
taran zard 'three pounds'	*ḫamas līrāt* 'five pounds'

Finally, while Domari lacks a definite article, the Arabic definite article *l-* is employed with definite noun phrases where both the noun and the numeral-attribute are derived from Arabic:

(32) mar-d-e l-ʔarbaʕ ḫurfān
kill-PFV-3PL DEF-four lamb.PL
'They slaughtered the four lambs.'

(33) dīr-os it-tānye eh-r-i muhandis-ēk
daughter-3SG DEF-second.F become-PFV-F engineer-PRED.SG.F
'Her other daughter became an engineer.'

4 Conclusion

The comparison with Macalister's (1914) materials offers some scope for observations in respect of the historical development of contact-induced change over the past century in at least two areas of structure, namely the loss of Turkish-derived vocabulary as well as of some of the inherited Indo-Aryan vocabulary (around 55 words are attested in Macalister's materials that were not familiar to the speakers I interviewed), and the adoption of fully-inflected modal and aspectual auxiliaries, compared to their use as impersonal forms in Macalister's material. One has to bear in mind, however, that Macalister's corpus is based on work with just a single speaker. Nevertheless, these changes provide some indication that the impact of Arabic continued to expand during the last century in which the language was spoken, a period during which the Doms lost much of their distinct culture and lifestyle as a result of the shift from a semi-nomadic service economy to a settled, wage-based but still socially isolated and stigmatised community.

The impact of Arabic on Domari prompts a theoretical challenge around identifying a form of the language that is structurally inseparable from Arabic. This can be illustrated by the following two examples:

(34) a. Domari
aktar min talātīn ḫamsa w talātīn sana mā lak-ed-om-is
more from thirty five and thirty year NEG see-PFV-1SG-3SG.OBL

b. Arabic
aktar min talātīn ḫamsa w talatīn sana mā šuft-ha
more from thirty five and thirty year NEG see.PRF.1SG-3SG.F
'I haven't seen her for more than thirty, thirty five years.'

(35) a. Domari
kān ʕumr-om yimken sitte snīn sabʕa snīn
be.PRF.3SG.M age-1SG maybe six years seven years

b. Arabic
kān ʕumr-i yimken sitte snīn sabʕa snīn
be.PRF.3SG.M age-OBL.1SG maybe six years seven years
'I was maybe six or seven years old.'

Both (34a) and (35a) are unambiguously identifiable to speakers as Domari utterances; moreover, their meaning cannot be conveyed in Domari in any other

way. Yet they each differ in just one single element from their respective counterpart Arabic utterances in (34b) and (35b): the use of the lexical verb with subject and object agreement (Domari *lak-ed-om-is* 'I saw her', Arabic *šuf-t-ha*) in the first, and the use of the 1SG possessive marker (Domari *-om*, Arabic *-i*) with the word *ʕumr* 'age' in the second. Despite being isolated examples, (34)–(35) illustrate the considerable extent of structural overlap between the two languages. Furthermore, the examples discussed above of bilingual suppletion in number agreement and adjective comparison, and the productive use of Arabic person agreement inflection with auxiliaries and with some complementisers and secondary object markers, mean in effect that active command of Arabic is a pre-requisite for speaking Domari.

It follows that Domari provides us with an opportunity to reconsider the taxonomy of contact-induced language change phenomena. It is not a Mixed Language by conventional definitions (cf. Bakker & Matras 2013; Matras 2009: chapter 10) since the Indo-Aryan source of grammatical inflection in all word classes is overwhelmingly consistent with the source of basic lexical vocabulary and of deictic and anaphoric elements (demonstrative and personal pronouns, interrogatives, and spatial adverbs). Impressionistically speaking, it is a language with "heavy borrowing" in that it shows the adoption of Arabic-derived material in a wide range of different structural categories. But the distribution of some of this material, taking into account the ubiquitous active bilingualism among Domari speakers, lends itself to the postulation of several particular types of contact-induced structural change, which I have labeled above fusion (wholesale non-separation of languages around a particular structural category, e.g. clause connectors and modal auxiliaries), inflectional compartmentalisaton (the use of Arabic inflectional paradigms with particular functional categories, notably modal and aspectual auxiliaries), and bilingual suppletion (activation of speakers' full command of Arabic vocabulary and inflection for creative formations around number agreement and adjective comparison).

Further Reading

Matras (2007) outlines contact influences on Jerusalem Domari in the context of a collection of chapters on contact-induced change in a sample of different languages.

Matras (2012) provides a descriptive and historical overview of Jerusalem Domari and includes extensive discussion of contact-induced change in the individual chapters as well as a chapter devoted to the impact of Arabic.

Matras (2009) is a general theoretical discussion of contact-induced change in functional-typological perspective and includes many examples from Jerusalem Domari.
Finally, Matras (2015) discusses patterns of morphological borrowing and their theoretical implications and gives as one of the examples the compartmentalisaton of modal and aspectual auxiliaries in Jerusalem Domari.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person
ABL	ablative
BEN	benefactive
COMP	complementiser
DAT	dative
F	feminine
IMP	imperative
IMPF	imperfect (prefix conjugation)
IND	indicative
ITR	intransitive
LOC	locative
M	masculine
OBL	oblique
PFV	perfective
PRED	predication (non-verbal)
PRS	present
PRF	perfect (suffix conjugation)
PROG	progressive
PST	past
RECP	reciprocal
REFL	reflexive
REL	relativiser
SBJV	subjunctive
SG	singular
TR	transitive

References

Bakker, Peter & Yaron Matras (eds.). 2013. *The mixed language debate: Theoretical and empirical advances*. Berlin: Mouton de Gruyter.

Black, George Fraser. 1913. The gypsies of Armenia. *Journal of the Gypsy Lore Society (new series)* 6. 327–330.

Field, Frederick. 2002. *Linguistic borrowing in bilingual contexts*. Amsterdam: John Benjamins.

Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the world's languages: A comparative handbook*. Berlin: De Gruyter Mouton.

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.

Kruse, Friedrich. 1854. *Ulrich Jasper Seetzen's Reisen durch Syrien, Palästina, Phönicien, die Transjordan-Länder, Arabia Petraea und Unter-Aegypten*. Vol. 2. Berlin: Reimer.

Macalister, Robert Alexander Stewart. 1914. *The language of the Nawar or Zutt, the nomad smiths of Palestine*. Edinburgh: Edinburgh University Press.

Matras, Yaron. 1999. The state of present-day Domari in Jerusalem. *Mediterranean Language Review* 11. 1–58.

Matras, Yaron. 2000. Two Domari legends about the origin of the Doms. *Romani Studies* 10. 53–79.

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.

Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.

Matras, Yaron. 2011. Domari. In Tatiana I. Oranskaia, Yulia V. Mazurova, Andrej A. Kibrik, Leonid I. Kulikov & Aleksandr Y. Rusakov (eds.), *Languages of the world: New Indo-Aryan languages*, 775–811. Moscow: Academia.

Matras, Yaron. 2012. *A grammar of Domari*. Berlin: De Gruyter Mouton.

Matras, Yaron. 2015. Why morphological borrowing is dispreferred. In Francesco Gardani, Peter Arkadiev & Nino Amridze (eds.), *Borrowed morphology*, 47–80. Berlin: Mouton de Gruyter.

Moravcsik, Edith A. 1978. Universals of language contact. In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of human language*, vol. 1: *Method and theory*, 94–122. Stanford: Stanford University Press.

Newbold, F. R. S. 1856. The Gypsies of Egypt. *Journal of the Royal Asiatic Society of Great Britain and Ireland* 16. 285–312.

Paspati, Alexandre. 1870. *Etudes sur les Tschinghianés ou Bohémiens de l'Empire Ottoman*. Osnabrück: Biblio.

Patkanoff, K. P. 1907. Some words on the dialects of the Transcaucasian Gypsies. *Journal of the Gypsy Lore Society (new series)* 1. 229–257.

Pott, August F. 1907. Über die Sprache der Zigeuner in Syrien. *Zeitschrift für die Wissenschaft der Sprache* 1. 175–186.

Part III

Domains of contact-induced change across Arabic varieties

Chapter 10

New-dialect formation: The Amman dialect

Enam Al-Wer
University of Essex

One fascinating outcome of dialect contact is the formation of totally new dialects from scratch, using linguistic stock present in the input dialects, as well as creating new combinations of features, and new features not present in the original input varieties. This chapter traces the formation of one such case from Arabic, namely the dialect of Amman, within the framework of the variationist paradigm and the principles of new-dialect formation.

1 Contact and new-dialect formation

1.1 Background and principles

The emergence of new dialects is one of the possible outcomes of prolonged and frequent contact between speakers of mutually intelligible but distinct varieties. The best-known cases of varieties that emerged as a result of contact and mixture of linguistic elements from different dialectal stock are the so-called colonial varieties, namely those varieties of English, French, Spanish and Portuguese which emerged in the former colonies in the Southern Hemisphere and the Americas.[1] In addition to colonial situations, the establishment of new towns can also lead to the development of new dialects; a case in point is Milton Keynes (UK), which was investigated by Paul Kerswill.[2] For Arabic, similar situations of contact are abundant, largely due to voluntary or forced displacement of populations,

[1]Among the studies that investigated such varieties are: Trudgill (2004), Gordon et al. (2004), Sudbury (2000) and Schreier (2003) for English; Poirier (1994; cited in Trudgill 2004) for French; Lipski (1994) and Penny (2000) for Spanish; and Mattoso (1972) for Portuguese.

[2]See Kerswill & Williams (2005).

Enam Al-Wer. 2020. New-dialect formation: The Amman dialect. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 231–246. Berlin: Language Science Press. DOI:??

growth of existing cities and the establishment of new ones. To date, however, the only study of a brand new dialect is the on-going investigation of the dialect of Amman, the capital city of Jordan, which is anticipated to provide a model for the study of dialect contact and koinéization in other burgeoning conurbations elsewhere in the Arab World. The bulk of this chapter will be dedicated to the details of this case.

Several other studies in Arab cities have focused on contact as a primary agency through which innovations permeate in the speech of migrant groups. Although no new dialects emerge in such situations, new patterns and inter-dialectal forms are common. For instance, Al-Essa (2009) reports that among the residents of the city of Jeddah, those who originally emigrated from various locations in Najd generally converge to the dialect of Jeddah, but also use innovations that do not occur in the target dialect, such as the second person singular feminine suffix *-ki* in words ending in a consonant, as in *ʔumm-ki* for Najdi *ʔumm-its* and Jeddah *ʔumm-ik* 'your (SG.F) mother'. Similarly, Alghamdi (2014) found interdialectal forms of the diphthongs /ay/ and /aw/ (viz. narrow diphthongal variants [ɛi], [ɔʊ]), as well as the monophthongs [ɛː] and [ɔː], in the speech of Ghamdi migrants who originally came to Mecca from Al-Bāḥa in the southwest of Saudi Arabia. In Casablanca, rapid urbanization led to immigration of large numbers of groups from all over Morocco, and subsequent contact between different dialects. Hachimi (2007: 97) suggests that this situation resulted in "the disruption of the rural/urban dichotomy that once dominated Moroccan dialects and identities", and the emergence of new categories of identification, which are symbolized through the usage of a mixture of features from different dialects.

In this context, it is worth pointing out some methodological challenges concerning the measurement of contact as an independent variable in quantitative sociolinguistics, and some improvements that have been made in research on Arabic. Contact is often invoked as an explanatory factor in contact linguistics in general, and has indeed been incorporated in theoretical formulations (e.g. Thomason & Kaufman 1988). In quantitative sociolinguistics, however, analysis of contact as a constraint on linguistic variation requires treating it as a variable from the outset of research, and finding ways to quantify it, in esssentially the same way that social categories such as age, gender and class are factored into the analysis. But how can contact be quantified? Recognizing the crucial role that (dialect) contact plays in the structure of variation and mechanisms of change, a number of quantitative studies have tested various methods of quantification. Al-Essa's (2009) study, mentioned above, was the first known quan-

tification of contact in studies of this sort. In order to do this, she measured the speakers' level of exposure to the target features through an index, consisting of a four-point scale, which gave a numerical value to each speaker's level of contact. Four criteria were used to determine the numerical value assigned to each speaker: friendships at school and work; involvement in neighbourhood affairs; friendship with speakers of the target dialect; kinship and intermarriage in the family (Al-Essa 2009: 208). Alghamdi's (2014) study in Mecca utilized and adapted Chambers' (2000) concept of regionality, by devising a regionality index based on the speakers' date of arrival in the city and place of residence. In Al-Wer (2002a), I suggested that in some cases level of education may be treated as an indication of level of contact with outside communities; and Horesh (2014) elicited information that was indicative of levels of contact between the speakers' L1 Arabic, and L2 Hebrew, which were later converted into factor groups, one of which was language of education, thus demonstrating that type of education can also be used to measure contact.[3]

1.2 Theoretical framework

The study of the formation of new dialects is credited particularly to the work of Peter Trudgill. In his *Dialects in contact* (1986) he laid the theoretical foundations of research in the field, arguing that 'face-to-face interaction' is a prerequisite for linguistic adaptation and diffusion of linguistic innovations.[4] Focusing on the formation of New Zealand English, Trudgill (2004) suggests a three-stage approach to dialect formation, which roughly corresponds to three successive generations of speakers.[5] These stages are very briefly summarized below, and illustrated using examples from Amman in §2.3.[6]

> Stage I (first generation): rudimentary leveling. This stage stipulates that at the initial point of contact and interaction between adult speakers of different regional and social varieties, minority and very localized linguistic features are leveled out.

[3]Several additional doctoral theses, currently in preparation at the University of Essex, address this issue.

[4]Trudgill (1986) integrated insights from Accommodation Theory (Giles 1973) in the study of dialect contact.

[5]In the same year, and based on the same data, the team working on the Origins of New Zealand English (ONZE) project, in which Peter Trudgill participated, also published a co-authored book on the topic (see Gordon et al. 2004).

[6]Trudgill (2004: 83–128) discusses and illustrates each stage with data from the ONZE corpus.

> Stage II (second generation): variability and mixing. At this stage, the first locally-born generation of children are presented with a plethora of features to choose from. Their speech contains considerable inter-individual and intra-individual variability, and new combinations of features.
>
> Stage III (third generation): emergence of stable and relatively uniform dialect. At this stage, focusing (Le Page & Tabouret-Keller 1985; see §1.3.2) gives rise to a crystalized dialect.

Trudgill (2004: 149)[7] concludes that the processes of dialect mixture and new-dialect formation are not haphazard but "deterministic in nature", "mechanical and inevitable"; and that, in *tabula rasa* situations, social and attitudinal factors do not play a role in the formation of new dialects.[8] "Determinism" in new-dialect formation and "the minor role that social factors, such as identity, play in *tabula rasa* situations" have instigated a wide and interesting debate among scholars. For instance, Tuten (2008: 261) proposes that "community identity formation and koiné formation are simultaneous and mutually dependent processes". Mufwene (2008: 258) agrees that common identity "is not part of the processes that produce new dialects"; but rather a by-product of it. Schneider (2008) elaborates on two issues: the relationship between accommodation and identity, and "the changing role of identity" in different colonial and postcolonial phases (2008: 262), pointing out cases of features from colonial varieties where the origins and spread of these features coincided with "a heightened national or social awareness" (2008: 266). Bauer (2008) contests Trudgill's implicit suggestion that accommodation leads directly to dialect mixing, on the basis that individuals vary in the extent to which they accommodate to others, and vary depending on the context; and in some cases no accommodation takes place, that is, accommodation is sporadic. He maintains that "it is not the accommodation as such that leads to dialect mixing; rather, it is the use that accommodation is put to by the next generation that leads to dialect mixing" (2008: 272). On the role of identity, Bauer contends that the very choice of a particular variant over another is indirectly an expression of "complex kinds of identity" (2008: 273).[9]

[7] See also Trudgill et al. (2000) and Trudgill (2008).

[8] Cf. Labov's (2001) principle of density.

[9] For more details, see Bauer (2008); and for Trudgill's responses to these points, see the discussion and rejoinder in *Language in Society,* 2008, vol. 37.

1.3 Mechanisms

The mechanisms involved in new-dialect formation fall under two broad headings: *koinéization* and *focusing*. Below are brief explanations of these mechanisms, to be followed by illustrations from data from Amman in the relevant sections.

1.3.1 Koinéization

Trudgill (2004: 84–88) uses *koinéization* as an umbrella term to refer to five processes, which operate at the same or different stages in the formation of new dialects: (i) mixing, which, as the name suggests, involves the use of features which originally came from different dialects; (ii) leveling, which involves gradual reduction and ultimate loss of minority features, that is, features that have least representation in the dialect mix; (iii) unmarking, a sub-type of leveling, which refers to the survival of unmarked and more regular forms even if they are not the majority forms; (iv) interdialect development, which are forms that arise out of interaction between different forms in the original mix, and can include phonetically, morphologically and syntactically intermediate forms; (v) reallocation, which refers to the survival of more than one variant of the same feature, which then undergoes reallocation in the new system; reallocation can be linguistic, social or stylistic.

1.3.2 Focusing

This term was introduced into sociolinguistics by Le Page & Tabouret-Keller (1985) to refer to the process whereby the new system "acquires norms and stability". A focused dialect contrasts with a diffuse (or non-focused) linguistic situation, where there is no consensus over norms, and no stability of usage.[10]

2 Dialect formation in Amman

2.1 History and demographics

Amman has no traditional dialect simply because until relatively recently it had no indigenous inhabitants. Though an important centre in ancient times, it remained largely deserted until the early years of the twentieth century.[11]

[10] See Le Page & Tabouret-Keller (1985: 181–182).

[11] Amman's ancient history is traced to the Ammonites (eighth century BCE), who called it *Rabbath Ammon* 'the great (or royal) city of the Ammonites'; the Romans changed its name to

In 1921, it was designated as the capital of Transjordan (the land east of the River Jordan), which became the Kingdom of Jordan in 1946. It thus attracted migrants from other parts of the country, as well as from Palestine, Syria and Lebanon. By the 1930s, the population had grown to 10,000 inhabitants, and by 1946 it stood at 65,000. The early migrants consisted of two groups: (i) the majority were economic migrants (traders and shop keepers as well as labourers) or civil servants, who were appointed in the state administration; (ii) the rest were political activists (mostly individuals from Syria and Lebanon, which were then still under French colonialist rule). The first group included families from both sides of the River Jordan, namely indigenous Jordanians from the east side, and Palestinians from all parts of historical Palestine. Statistics regarding numbers from each group are unavailable, but I was able to collect fairly reliable information, through ethnographic interviews, about the provenance of a large sector of the first generation of migrants. According to my research, the vast majority came from two particular locations: the Jordanian city of Sult (20 kilometres northwest of Amman), and the Palestinian city of Nablus (110 kilometres from Amman).

The city continued to receive waves of migrants from other locations in Jordan and from Palestine, especially following the two wars in 1948 and 1967, which resulted in the occupation of historical Palestine, and the displacement of well over three million Palestinians over the years, most of whom sought refuge in Jordan. Between 1950 and 1990, the population of Amman doubled more than fifteen times, to reach approximately two million by 2004. According to the 2018 census estimate, the city is home to 2,554,923 Jordanian nationals, and 1,452,603 non-Jordanians, that is, a total of over four million people live in the city currently.[12] Given the political situation in the region, the population of Amman is forecast to reach six million by 2025.

Against this demographic background, there are three important points to note:

1. There is no geographically neutral variety of spoken Jordanian Arabic. All speakers therefore use some form of local dialect, regardless of social class.

2. Whereas in neighbouring countries (Syria, Egypt, Lebanon, Palestine), the dialect of the capital acts as a standard prestigious norm, Jordan never had

Philadelphia; the Arab Ummayads took over in the seventh century CE and restored its Semitic name, *Amman.*

[12]Department of Statistics, Jordan: http://dosweb.dos.gov.jo/DataBank/Population_Estimares/PopulationEstimates.pdf (accessed 06/01/2020).

a linguistic centre of its own.

3. Jordanians, and Palestinians, generally identify themselves with the area in which their forebears lived, rather than the locality in which they were born and bred. However, recently a growing number of inhabitants of Amman (particularly third and fourth generations) have begun to identify themselves as 'Ammanis'.

The emergence of a distinctive and focused dialect of Amman, in tandem with the emerging Ammani identity, represents a radical shift in the sociolinguistic patterns from a plethora of local varieties to a situation similar to that described for neighbouring states in §1.

2.2 The Amman Project

This research traces the formation of this new dialect from inception to stabilization over three generations, spanning a period of approximately the last eighty years. It initially focused on generational differences, by investigating the developments in the speech of three generations of families who originally came from the Jordanian city of Sult and the Palestinian city of Nablus; this initial investigation confirmed the following hypotheses:

> A new dialect has emerged and its usage has stabilized.
>
> This dialect is unique – it grew as an outcome of the contact between Jordanian and urban Palestinian dialects, but is distinct from these input varieties.
>
> The formation of a distinctive dialect in Amman is closely associated with relative stabilization in the population, possibly during 1970–1990, and the development of an Ammani community with its own identity.[13]

The second phase of the research focused on the younger generation from affluent West Amman; and the final phase, ongoing, expands the sample to include speakers from less affluent East Amman. Altogether, the research aims to collect data from approximately 120 speakers, from both sides of the city. The project on Amman itself is complemented by past and ongoing research (by myself and others) on sociolinguistic trends in areas outside Amman, which provide two valuable types of relevant information: (i) further evidence of the input varieties;

[13] Full details can be found in Al-Wer (2002a; 20030).

(ii) spreading of innovative features of the Amman dialect to other parts of the country.

The framework of analysis adopted here is the Variationist Sociolinguistic Paradigm, as described in Labov's trilogy (1994; 2001; 2010). More specifically, the project is guided by the principles of dialect contact and new-dialect formation, as outlined in §1. As discussed earlier, one of the dominant issues in the study of the formation of new dialects is the debate over the types of factors which determine it. The Amman project offers an opportunity to investigate these issues in detail, particularly because it is still possible to trace the different stages of formation over the three generations of native inhabitants.

2.3 Formation over three generations

Based on the analysis of speech samples from three generations of Ammanis, the formation of the dialect is a textbook case. Many of the processes of koinéization explained above are operative, as will be demonstrated presently.

2.3.1 Stage I: first generation

The first generation arrived in Amman during the 1930s as adults. The most noticeable aspect of their speech is that it can easily be identified with the original dialects of the places from which they migrated, while localized features are leveled out (cf. rudimentary leveling; Trudgill 2004). The features which are lost at this stage are summarized below.

> *Jordanian input.* Traditional Jordanian dialects, including the dialect of Sult, are known to affricate /k/ to [ʧ] in front-vowel environments generally, as in /keːf/ > [ʧeːf] > 'how'. This feature is still widely used, especially in northern varieties, as well as in Sult[14] – where most of the early migrants in Amman came from. Already in the first generation, this feature is completely lost; all instances of this variable were rendered with [k]. In other words, first-generation speakers deaffricate /k/.[15] Although conditional affrication of /k/ is fairly widespread in the region's rural dialects, and is certainly not a minority feature in Jordanian dialects, its use is heavily stigmatized, and none of the urban dialects have it. Stigmatization is the likely reason that motivates the loss of this feature.

[14]On this feature in Traditional Jordanian dialects, see Al-Hawamdeh (2015), Herin (2010) and Herin & Al-Wer (2013).

[15]See Al-Wer (2007) for more details.

Also characteristic of the traditional dialects is the maintenance of a gender distinction in the second and third person plural pronouns, and pronominal, verbal and nominal suffixes. For example: *ʔintu* 'you (PL.M)', *ʔintin* 'you (PL.F)'; *ʔumm-hum* 'their (M) mother', *ʔumm-hin* 'their (F) mother'; *rāḥu* 'they (M) have left', *rāḥin* 'they (F) have left'; *ḥilwīn* 'pretty (PL.M)', *ḥilwāt* 'pretty (PL.F)'. What we find in Amman is gender neutralization in these forms, such that the masculine form is used to refer to both genders. The traditional system is currently variable in all major Jordanian cities, and seems to be giving way to a neutralized form, as in Amman, which is an indication that Amman has become a focal point from which linguistic innovations radiate. No particular social value is attached to the traditional feature (maintenance of gender distinction) although there is awareness that it is characteristically found in provincial towns and villages. The observation that it has become variable in many cities and towns means that it is also becoming a minority feature in urban areas in particular. The change affecting this feature in Jordanian dialects in general may be described as a form of simplification, where the number of distinct forms in the paradigm as a whole is reduced. Additionally, none of the urban Palestinian dialects maintain a gender distinction in these categories. In contact situations especially, the direction of change is normally towards the simpler system (the "Simplification Preference"; Lass 1997: 253).[16]

Palestinian input. In urban Palestinian, the high-frequency terms *mbāriḥ* 'yesterday' and *sāʕa* 'hour/time' are pronounced with raised vowels: *mbēriḥ* and *sēʕa.* This is an extremely marked pronunciation in the context of Jordan, as no Jordanian dialect has it. It is also a feature that is overtly commented upon, and often used to mimic dialects that have it. Extreme raising of /ā/ generally is a hallmark of many urban Palestinian dialects, most notably in the dialect of Jerusalem; as will be explained later, third-generation speakers with urban Palestinian heritage use considerably lower variants than first- and second-generation speakers from the same group. It is possible that lowering in these high-frequency items in the first generation is the onset of the change that escalated in successive generations. In the third generation of this group, the speakers change their pronunciation in these items only, but continue to use noticeably higher variants of /ā/

[16]It should be pointed out that the urban Palestinian dialect, similarly to all city dialects in the region, has the *-on and -kon* endings, which are used with both genders, rather than masculine *-um* and *-ku*, which are the koiné forms in modern Jordanian dialects; see Al-Wer (2003) for more details about this feature.

in other items, for example, 'Amman' is pronounced as [ʕəmmɛːn]; *fālit* 'loose' is pronounced as [fɛːlɪt].

2.3.2 Stage II: second generation

This is the first locally-born generation; the majority of the speakers in the sample fall in this category, or arrived as very young children (under ten). The speech of members of this generation shows extreme inter-speaker and intra-speaker variability, and a mixture of features from both norms (Jordanian and urban Palestinian). For example, the same speaker is found to use the second person plural pronominal suffixes *-ku* (Jordanian) and *-kon* (Palestinian), e.g. *kēf ḥāl-ku ~ kīf ḥāl-kon* 'how are you (PL)'. In this example, we also find alternation in the vowel of the item *kēf ~kīf* 'how'; the former is Jordanian while the latter is typical of urban Palestinian (and urban Levantine in general). The data also contained a mixture of Jordanian and Palestinian third person plural suffixes *-hum* and *-hon,* e.g. *šift-hum ~ šift-hon* 'I have seen them'. At the level of phonology, speakers in this generation use a mixture of Jordanian [g] and urban Palestinian [ʔ], which are variants of historical /q/; and a mixture of interdental and stop counterparts of /θ/, /ð/ and /ð̣/. Importantly, in this generation there is a complication in sociolinguistic correlations: whereas in the first generation there is a one-to-one relationship between origin and the dialect used, in the second generation certain groups from both backgrounds use features characteristic of the other group's dialect. The particular sub-groups that do this are the Jordanian women, who in this generation use Palestinian [ʔ] almost consistently, as well as a high rate of the stop variants of interdentals (see above), and use both Jordanian *ʔiḥna* and Palestinian *niḥna* 'we'. The second most divergent group (from their heritage dialect) is Palestinian men; they use Jordanian [g] at a rate of 50%, or more in some cases. The remaining groups, Jordanian men and Palestinian women, are considerably more conservative with respect to their heritage variants, although they too are variable. What this pattern shows is that gender emerges as an important social factor in this generation, in addition to dialectal heritage, which continues to influence individuals' behaviour, but interacts with gender at the same time.

2.3.3 Stage III: third generation

Third-generation Ammanis were all born in the city (in the 1970s). They diverge from their parents' and grandparents' dialects, and speak a clearly distinct dialect, regardless of their own dialectal heritage. The mixture and variability we saw in

the second generation is much reduced in the third generation; there is, instead, stability in the usage of many features, including intermediate fudged forms, new patterns, and new features that were not present in the input varieties. The third generation agree on the characteristics of Ammani, and have intuitions as to what you can and cannot say in this dialect. Importantly, they express affiliation with the city; for instance, they identify themselves as 'Ammanis', by which they mean that they are native to the city. In other words, the formation of the dialect is simultaneously a formation of a community.

In this generation, gender emerges as a major organizing category; for instance, all the women in this generation, regardless of dialectal heritage, use [ʔ] consistently, while the men continue to use both [g] and [ʔ]. The variability in men's speech is constrained by context and interlocutor for the most part; whereas the speech of women is not subject to these constraints.[17] The development in the use of variants of the variable (q), as explained above, is a clear example of a variable that has undergone social and stylistic reallocation (see §1 above) in the sense that both variants [g] and [ʔ], which originally come from different dialects in the input varieties, have survived the koinéization process but no longer signify ethnicity or dialectal background straightforwardly; the use of one or the other is now subject to layers of constraints. As far as the interdental sounds are concerned, both gender groups use the stop variants more often than the interdental variants. But while the men vary between affricate [ʤ] and fricative [ʒ] of the variable (ǧ), the women use [ʒ] almost consistently.

In addition to the features listed under stage III, the following features are at an advanced stage of focusing in the new dialect:

> The feminine ending *-a*. The input varieties differ in the phonology and phonetics of the realization of the feminine ending in the unbound state. In traditional Jordanian, the low vowel [a] is the default choice, except after coronal sounds, where it is raised to [ɛ]. In urban Palestinian dialects, the default choice is [e], or raised [ẹ], except after pharyngeal and emphatic consonants in general.[18] In Amman, the third generation consistently use a fudged form made up of urban Palestinian phonology and the Jordanian phonetic property of the raised vowel, such that they raise /a/ to [ɛ] except after back sounds, e.g. *mukinsɛ* 'broom', *ḥilwɛ* 'pretty', *ṣaʕbɛ* 'difficult', but *rāyḥa* 'has gone/is going'; *žāmʕa* 'university'; *ṣulṭa* 'authority'.

[17] Details about this feature can be found in Al-Wer & Herin (2011).

[18] A preceding /r/ blocks raising in general unless there is an /i/-type vowel in the environment; for a complete account of the phonology of the feminine ending, see Al-Wer et al. (2015).

In morphophonology, a new form of the second person plural suffix has emerged, and is used consistently. The input forms are: Jordanian *-ku*, as in *gultilku* 'I have told you (PL)'; and urban Palestinian *-kon,* as in *ʕmiltilkon* 'I have made for you (PL)'. The form that has been focused in Amman is *-kum,* thus *ʕindkum* 'you (PL) have'; *ḥakētilkum* 'I have told you (PL)'. The success of this form, rather than one from the input varieties, is explained with reference to markedness and simplification.[19]

In morphology, the input varieties differ in the conjugation of the third person masculine imperfect verb form. In both dialects, Jordanian and Palestinian, the imperfect takes a *b-* prefix, but whereas in Jordanian dialects *yod* is dropped from the stem in the *b*-imperfect in all environments, in Palestinian dialects it is dropped in open syllables only. For example: Jordanian *biḥki* 'he talks', *binuṭṭu* 'they jump'; urban Palestinian *byiḥki, binuṭṭu.* Ammanis (third generation) drop *yod* everywhere except where it carries person information, namely in glottal-initial verbs *ʔakal* 'to eat', and *ʔaxað* 'to take'; thus we get *biḥki, binuṭṭu,* but *byākul* 'he eats' (stem *ʔakal* 'to eat'), *byāḫdu* 'they take' (stem *ʔaḫað* 'to take').[20]

3 Conclusion

The formation of the Amman dialect is simultaneously the formation of a community; and the social factors involved in the formation of the dialect evolve and realign accordingly. One of the most interesting aspects of this process is that none of the factors become totally irrelevant. For instance, dialectal heritage – which, in the case in hand, coincides with ethnicity (Jordanian/Palestinian) – is the most important predictor in the speech of the first generation. In the second generation, gender emerges as an important factor, but the linguistic developments at this stage can only be understood as an interaction between the old and new social constraints; for instance, in stage II, it is not merely women who use [ʔ] rather than [g], but it is *Jordanian* women who diverge from their heritage variant; and it is not the behaviour of men in general that explains the evolution in the re-distribution of these variants, but specifically the behaviour *Palestinian* men. These two sub-groups (Jordanian women and Palestinian men) are responsible for the diversification of, firstly, their respective group's linguis-

[19]For analysis of this development, see the full details in Al-Wer (2003).

[20]There are further complications and variations in the conjugations of these verbs; for these details see Al-Wer (2014).

tic repertoire and consequently the repertoire of the linguistic system that is passed on to the next generation. In stage III, the third generation's behaviour responds to two riders: the system inherited from their parents and the changes in the socio-political environment around them. A further realignment of social factors occurs in response, and new constraints are added to the old pile; at this stage, the inherited identifications of the variants involved – that is, [g] is Jordanian and appropriate for men, [ʔ] is Palestinian and appropriate for women – are reformulated through the addition of further new constraints, namely context and interlocutors. Consequently, the usage of the variants involved is redistributed according to style,[21] they acquire additional identifications and social meanings, and the social constraints are realigned, such that the role of ethnicity becomes subsidiary, while gender and style are the major organizing factors. The younger generation now define [ʔ] as "Ammani", and [g] as "authentic Jordanian". The meaning of "Jordanian" itself is often negotiated and expanded beyond the limits of ethnicity to denote a regional identity, recognizing citizenship as the primary defining component of membership in this group, although the old meaning (those whose roots lie on the east side of the river) is not obliterated altogether.[22] A further realignment of social factors in Amman involves type of profession, which is emerging as a constraint. This may have been precipitated by the expansion of the private sector over the past two decades or so, especially banking and the service industry in general, and the tourism industry. According to preliminary analysis of recently collected data, different types of employment, within and across the two sectors, fall within the realms of different linguistic markets.

The context in which the Amman dialect was formed was *tabula rasa* in the sense that there was no pre-existing Amman dialect. The obvious difference from, say the *tabula rasa* colonial situations, is that the early settlers in Amman were not isolated from their original communities or from Arabic speakers in the surrounding areas; social factors definitely play a role in the formation of the dialect in this case. The question therefore is not whether social and attitudinal factors are involved, but rather which social factors, how they evolved, and their relative importance.

[21]Style as a correlate of linguistic usage can mean different things; here I use it to refer to context (as in Labov 1972), and audience or interlocutor (as in Bell 1984). For details of how style evolved a sociolinguistic correlate, see Eckert & Rickford (2001).

[22]The question of "who is Jordanian" is, for many, a sensitive issue, which has often caused heated debates on various media platforms.

Further reading

Al-Wer (2011) provides a brief description of the Amman dialect.
Al-Wer (2007) summarizes the processes and dynamics of the formation of the dialect of Amman, along with a list of thirteen linguistic features that have been focused in this dialect.
Al-Wer (2002b) focuses on the long vowels and the realization of the feminine ending in the newly formed dialect of Amman.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person
F	feminine
L1	first language
L2	second language
M	masculine
ONZE	Origins of New Zealand English project
PL	plural
SG	singular

References

Al-Essa, Aziza. 2009. When Najd meets Hijaz: Dialect contact in Jeddah. In Enam Al-Wer & Rudolf De Jong (eds.), *Arabic dialectology: In honour of Clive Holes on the occasion of his sixtieth birthday*, 203–222. Leiden: Brill.

Al-Hawamdeh, Areej. 2015. *A sociolinguistic investigation of two Horani features in Souf, Jordan*. Colchester: University of Essex. (Doctoral dissertation).

Al-Wer, Enam. 2002a. Education as a speaker variable. In Aleya Rouchdy (ed.), *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme*, 41–53. London: Routledge.

Al-Wer, Enam. 2002b. Jordanian and Palestinian dialects in contact: Vowel raising in Amman. In Mari Jones & Edith Esch (eds.), *Language change: The interplay of internal, external and extra-linguistic factors*, 63–79. Berlin: Mouton de Gruyter.

Al-Wer, Enam. 2003. New dialect formation: The focusing of -kum in Amman. In David Britain & Jenny Cheshire (eds.), *Social dialectology: In honour of Peter Trudgill*, 59–67. Amsterdam: Benjamins.

Al-Wer, Enam. 2007. The formation of the dialect of Amman: From chaos to order. In Catherine Miller, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.), *Arabic in the city: Issues in dialect contact and language variation*, 55–76. London: Routledge.

Al-Wer, Enam. 2011. Amman arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*. Leiden: Brill.

Al-Wer, Enam. 2014. Yod-dropping in *b*-imperfect forms in Amman. In Reem Khamis-Dakwar & Karen Froud (eds.), *Perspectives on Arabic Linguistics XXVI: Papers from the Annual Symposium on Arabic Linguistics, New York, 2012*, 29–44. Amsterdam: 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.

Alghamdi, Najla Manie. 2014. *A sociolinguistic study of dialect contact in Arabia: Ghamdi immigrants in Mecca*. Colchester: University of Essex. (Doctoral dissertation).

Bauer, Laurie. 2008. A question of identity. *Language in Society* 37(2). 270–273.

Bell, Alan. 1984. Language style as audience design. *Language in Society* 13(2). 145–204.

Chambers, Jack. 2000. Region and language variation. *English World-Wide* 21. 1–31.

Eckert, Penelope & John R. Rickford. 2001. *Style and sociolinguistic variation*. Cambridge: Cambridge University Press.

Giles, Howard. 1973. Accent mobility: A model and some data. *Anthropological Linguistics* 15(2). 87–105.

Gordon, Elizabeth, Lyle Campbell, Jennifer Hay, Margaret MacLagan, Andrea Sudbury & Peter Trudgill. 2004. *New Zealand English: Its origin and evolution*. Cambridge: Cambridge University Press.

Hachimi, Atiqa. 2007. Becoming Casablancan: 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.

Herin, Bruno. 2010. *Le parler arabe de Salt (Jordanie)*. Brussels: Université Libre de Bruxelles. (Doctoral dissertation).

Herin, Bruno & Enam Al-Wer. 2013. From phonological variation to grammatical change: Depalatalisation of /č/ in Salti. In Clive Holes & Rudolf De Jong (eds.),

Ingham of Arabia: A collection of articles presented as a tribute to the career of Bruce Ingham, 55–73. Leiden: Brill.

Horesh, Uri. 2014. *Phonological outcomes of language contact in the Palestinian Arabic dialect of Jaffa*. Colchester: University of Essex. (Doctoral dissertation).

Kerswill, Paul & Ann Williams. 2005. New towns and koineization: Linguistic and social correlates. *Linguistics* 43. 1023–1048.

Labov, William. 1972. *Sociolinguistic Patterns*. Oxford: Blackwell.

Labov, William. 1994. *Principles of linguistic change*. Vol. 1: *Internal factors*. Oxford: Blackwell.

Labov, William. 2001. *Principles of linguistic change*. Vol. 2: *Social factors*. Chichester: Wiley-Blackwell.

Labov, William. 2010. *Principles of linguistic change*. Vol. 3: *Cognitive and cultural factors*. Chichester: Wiley-Blackwell.

Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.

Le Page, R. B. & Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.

Lipski, John. 1994. *Latin American Spanish*. Harlow: Longman.

Mattoso, Joaquim. 1972. *The Portuguese language*. Chicago: University of Chicago Press.

Mufwene, Salikoko S. 2008. Colonization, population contacts, and the emergence of new language varieties: A response to Peter Trudgill. *Language in Society* 37(2). 254–258.

Penny, Ralph. 2000. *Variation and change in Spanish*. Cambridge: Cambridge University Press.

Poirier, Claude. 1994. La langue parlée en Nouvelle-France: Vers une convergence des explications. In Raymond Mougeon & Édouard Beniak (eds.), *Les origines du français québécois*, 237–273. Sainte-Foy, Québec: Presses de l'Université Laval.

Schneider, Edgar W. 2008. Accommodation versus identity? A response to Peter Trudgill. *Language in Society* 37(2). 262–267.

Schreier, Daniel. 2003. *Isolation and language change: Contemporary and socio-historical evidence from Tristan da Cunha English*. London: Macmillan.

Sudbury, Andrea. 2000. *Dialect contact and koineisation in the Falkland Islands: Development of a Southern Hemisphere variety?* Colchester: University of Essex. (Doctoral dissertation).

Thomason, Sarah G. & Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.

Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Wiley-Blackwell.

Trudgill, Peter. 2004. *New-dialect formation: The inevitability of colonial Englishes*. Oxford: Oxford University Press.

Trudgill, Peter. 2008. Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation. *Language in Society* 37. 241–254.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis & Margaret Maclagan. 2000. Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36. 299–318.

Tuten, Donald. 2008. Identity formation and accommodation: Sequential and simultaneous relations. *Language in Society* 37(2). 259–262.

Arabic and contact-induced change

change blurb in localmetadata.tex