Grammatical borrowing in Khuzistani Arabic

Yaron Matras and Maryam Shabibi

1. Background

The Khuzistani dialect of Arabic (henceforth Kh. Arabic) is spoken natively by over 3 million people, who constitute roughly 7 percent of the population of the province of Khuzistan in south-western Iran. The dialect is the easternmost representative of the continuum of Mesopotamian dialects of Arabic, which cover the river lands of southern Iraq in the west (Ingham 1982: 1–4). Arab settlement in the area is believed to go back to the beginning of the Christian era. In the centuries following the advent of Islam, the Arabic language enjoyed the status of the literary language of religion, scholarship, and administration, as well as being the primary language of everyday communication in the province. This changed with the coming to power of Reza Shah Pahlavi in 1926 and the introduction of an intensive campaign favouring Persian as the only official state language. The policy included the settlement of a Persian-speaking population in the province.

Unlike other dialects of Arabic, Kh. Arabic has not attracted much attention within the linguistic community. Ingham (1997) devotes a chapter to the dialect, focusing, however, on an introductory discussion of phonology and vocabulary only; and Shabibi (2006) provides an overview of the structures of the dialect along with an analysis of contact-induced developments in morphosyntax.

Persian is now the only language of education, local media and newspapers, administration, and most urban commerce in the province of Khuzistan. Arabic is the language of the family and Arabic-speaking neighbourhoods, though even as an informal language it is now in decline, and Persian is the preferred language of the younger generation born since the 1970s. All educated adult speakers of Arabic are bilingual, and Arabic monolingualism is limited to the uneducated older generation, and to the older generation in rural communities. Arabic literacy is limited by and large to reading the Qur’an, and to a very basic level of instruction in Modern Standard Arabic, though even most educated Arabs have no active command of Modern Standard Arabic. There is, however, considerable exposure to Arabic-language satellite media, and so to the broadcast (oral) version of Modern Standard Arabic. In some
cases, Khuzistani Arabic speakers are able to read modern Arabic by drawing on their exposure to these media, combined with their basic familiarity with the Arabic script and with Classical Arabic (Qur’an).

2. Phonology

The only apparent phonological contact phenomenon in Kh. Arabic is the interchange of /ɣ/ and /q/, in words such as /ɣarîb, qarîb/ ‘close’ (cf. Modern Standard Arabic /qarîb/ ‘close’, /ɣarîb/ ‘strange’). This matches the realization in Persian of etymological /q/ as /ɣ/. Phonemes that are otherwise absent from the Arabic system, most notably /pl/, are retained in Persian loanwords: *panjara* ‘window’.

3. Morphological typology

A major change under Persian influence is the levelling of the status of attributes. In Arabic, adjectival attributes follow the head noun, and agree with the head noun in gender, number, as well as in definiteness:

(1) Standard Arabic (and other dialects)
   a. *walad kâbir*
      boy big.M
      ‘a big boy’
   b. *l-walad l-kâbir*
      DEF-boy DEF-big.M
      ‘the big boy’

Nominal attributes, by contrast, are conjoined by means of the attributive *Idâfa*-construction, whereby only the dependent (genitive) noun is overtly marked for definiteness:

(2) Standard Arabic (and other dialects)
    *walad l-mudîr*
    boy DEF-director
    ‘the director’s son’
In Persian, both types of attributes are treated in the same way: The attribute (whether adjectival or nominal) follows the head, and an attributive particle (the Ezāfe marker) mediates between the two:

(3) Persian
   a. *pesar-e bozorg*
      boy-EZ big
      ‘the big boy’
   b. *pesar-e modīr*
      boy-EZ director
      ‘the director’s son’
   c. *xūne-ye seftīd*
      house-EZ white
      ‘the white house’
   d. *moqālām-e madrese*
      teacher-EZ school
      ‘the school teacher’

The pattern in Kh. Arabic matches the Persian arrangement (note that, as in other dialects of Arabic, the definite article *l-* assimilates to dental consonants, resulting in gemination of that consonant):

(4) Khuzistani Arabic
   a. *walad ē-čibīr*
      boy DEF-big.M
      ‘the big boy’
   b. *walad l-modīr*
      boy DEF-director
      ‘the director’s son’
   c. *biāt l-abyad*
      house DEF-white
      ‘the white house’
   d. *moqālām-āt l-madrāsā*
      teacher-F.CONS DEF-school
      ‘the school teacher’

Note that in the adjectival attributive construction in (4a and 4c) overt definiteness agreement between noun and adjective is lacking, just like in the
genitive attribute construction in (4b and 4d). Based on the Persian model, Kh. Arabic has reanalysed the definite article in such constructions as a marker of attribution, which matches the Persian (definite) Ezāfe marker -(y)e. Its distribution now resembles that of the Persian Ezāfe attributive marker: It appears, like Persian -(y)e, between the two constituents of the attribution, and it is used to link both adjectival, and nominal attributes.

Further evidence that the functions of the Persian construction are mapped onto Arabic structures is provided by the position of the feminine Construct State or Idāfa-marker -at, seen in (4d) in a position that is not untypical of Arabic as a whole. In Arabic, the Construct State marker (still recognisable in the vernaculars only in the feminine singular) is reserved for nominal attribution, as in (4d). But in Kh. Arabic we find it in adjectival attributive constructions as well, as in (5a–b); it even attaches directly to adjectives, as in ṣaly-at ‘high.f’ in (5b):

(5)  Khuzistani Arabic
    a. jazīr-at l-xadra
        island-f.CONS DEF-green
        ‘the green island’
    b. tūf-at ṣaly-at l-biṣṭ
        wall-f.CONS high-f.CONS DEF-house
        ‘the high wall of the house’

This matches again the distribution of the Persian Ezāfe marker -(y)e (6):

(6)  Persian
    a. jazīre-ye sabz
        island-EZ green
        ‘the green island’
    b. divār-e boland-e xūne
        wall-EZ tall-EZ house
        ‘the high wall of the house’

Note that in the ‘mixed’ type, in (5b), involving both an adjectival-attribute (‘high wall’) and a genitive attribute (‘wall of the house’), the first (adjectival) attribution relies exclusively on the Construct State marker, while the second (nominal) relies on the combination of the Construct State marker with the following definite article. In fact, Kh. Arabic allows for variation in such cases, and the Construct State marker may be accompanied by a definite art-
icle in both positions. Consider example (7), where the nouns are masculine, and there is no option of using an overt Construct State marker:

(7)  a. Khuzistani Arabic
     \textit{walad ċ-čibir l-modir}
     boy DEF-big DEF-director
     'the director's big/eldest son'

     b. Standard Arabic (and other dialects)
     \textit{walad l-mudir l-kabir}
     boy DEF-director DEF-big
     'the director's big/eldest son'

     c. Persian
     \textit{pesar-e bozorg-e modir}
     boy-EZ big-EZ director
     'the director's big/eldest son'

The crucial aspect of the Kh. Arabic construction is (1) to have a marker of attribution mediating between the head and its attribute, (2) to place the attribute in a position immediately following its head, and (3) to avoid any overt marking of definiteness in the adjectival attribution. In all this, Kh. Arabic copies precisely the Persian attributive construction. Contrasting with Persian, it retains a distinct marking of attribution with feminine singulars, but allows this marking to assimilate into the generic function of the attributive marker. The outcome of the process is (1) the loss of the distinction between nominal and adjectival attribution, (2) the loss of overt marking of definiteness in attributive constructions, (3) a change in the word order in complex ('mixed') attributive constructions (as in 5b and 7a), and, finally, (4) gender variation in the marking of the attributive construction, with optional use of the definite article to accompany the Construct State in feminine singulars in complex attributions.

4. Nominal structures

The most notable contact-induced change in Kh. Arabic nominal structures is the status of the Idāfa-construction alluded to above. The replication of a construction type that is similar to the Persian Ezāfe leads, as discussed above, to the abandonment of definiteness agreement. The decline of overt definiteness marking can also be observed in other constructions in the language,
notably in the absence of a definite article with definite head nouns of relative clauses. This too follows a Persian model (where definiteness generally remains unmarked):

(8) Khuzistani Arabic

\textit{mara lli šift-ū-ha xābar-at.}

woman REL saw-2PL.M-3SG.F called-3SG.F

'The woman that you saw called.'

5. Verbal structures

In the derivation of verbs, the tendency to paraphrase inchoative and causative verbs drawing on an analytic construction rather than on derivational morphology, although found in other dialects of Arabic, appears to be reinforced by Persian. Thus we find:

(9) Khuzistani Arabic

\textit{š-šījra z-qyār-ā šwayye šwayy tšir čibir-ā.}

DEF-tree DEF-small-F little little become.3SG.F big-F

'The small tree gradually grows.'

Loan-verbs appear to be limited to the replication of Persian compound verbs consisting of a nominal stem (\textit{masdar}) and a verbalizing element or ‘light verb’ (Persian \textit{kardan ‘to do’} or \textit{šodan ‘to become’}). The nominal stem, often itself an Arabic loan into Persian, is replicated directly in Kh. Arabic, while as corresponding native light verb \textit{saww- ‘to do’} is employed for Persian \textit{kardan}, and \textit{gār- ‘to become’} for Persian \textit{šodan}, thus: Persian \textit{taayib-eš kard} ‘he followed him’ (follow-3SG did.3SG) is rendered \textit{sawwā-h taayib} (did.3SG-3SG follow).

An additional change to the verb system, brought about through Persian influence, concerns the tense system. Persian has both a simple past tense, which is expressed by the person-inflected past stem of the verb, and a composite past tense, which consists of a past participle and an auxiliary. The auxiliary, based on the existential verb, may inflect for person as well as tense; the present-tense auxiliary is used to form the perfect, the past-tense auxiliary forms the pluperfect. Arabic, by contrast, has only one, simple past tense, though combinations of the past-tense existential verb with the lexical verb (usually in the imperfect or present-future) are also possible, usually expressing habitual aspect of conditional mood. Kh. Arabic copies the Per-
sian composite past tense, drawing on inherited resources. The only available participle form in Arabic is the present participle, which inflects for gender and number (but not for person), and it is this form that serves as the basis for composite past tense in Kh. Arabic. Since the Arabic existential verb does not have a present-tense form, the only available auxiliary is a past-tense auxiliary; the construction thus matches the Persian pluperfect:

(10) a. Khuzistani Arabic
   
   *mān* *rāḥ-hat* *lā-l-bist₂* *huwā× mā-rāy-h* *čān*
   when went-1SG to-DEF-home he NEG-going.SG.M was.3SG.M
   ‘When I went home he had not gone away.’
   
   b. Persian
   
   *vāyi tech-tām* *xūn-e, ān na-rafe* *būd.*
   when went-1SG home he NEG-gone was.3SG.M
   ‘When I went home he had not gone away.’

(11) a. Khuzistani Arabic
   
   *mān* *gabul šāyf-at-ha* *čān.*
   from past seeing.SG.F-3SG.F was.1SG
   ‘I had seen her before.’
   
   b. Persian
   
   *az yābl ān-ō dīde būd-am.*
   from past 3SG-ACC seen was-1SG
   ‘I had seen her before.’

6. Other parts of speech

A series of Persian discourse markers, fillers, tags, and focus particles are used in Kh. Arabic. Most of these elements are well integrated into Kh. Arabic and are not perceived by speakers as foreign. The category that is most obviously influenced by Persian is that of discourse markers with a primarily interaction-qualifying rather than syntactic-semantic function: *xōb*xōb/xōs ‘well’, xōlāse ‘in sum’, abāte ‘of course’, ḥiḍ ‘at all, altogether’, ham ‘indeed, well’:

(12) *xōb* *w-hāy* *sabab ham* *l-laḍi gabal čān*
    
    *DM* and-this reason DM REL once was
    ‘Well, and this reason that indeed once existed for this …’
(13) *xolāse hīc mā-rah-na madrasa.*
DM DM NEG-went-1PL school
‘After all, we didn’t go to school at all.’

(14) *albate čān-an ham b-ōīc z-zamān banāt č-čān-an*
DM were-3PL.F DM in-that DEF-time girls REL-were-3PL.F
yarh-an.
go-3PL.F
‘Of course there were indeed girls at the time who used to go [to school].’

These are accompanied by Persian-derived focus particles: *ham* ‘too’ and *ham* ...
... *ham* ‘both ... and’.

(15) *ōīc Šond-ha Šnīn frāx ana ham Šond-i Šnīn.*
that.F poss-F two children I too poss-2SG two
‘She has two children, and I have two children, too.’

(16) *ham ana w ham Šali rāh-na l-šal-park.*
both I and also Ali went-1PL to-DEF-park
‘Both Ali and I went to the park.’

(17) *umm-i ham yāšl-at l-š-māṣīn ham naddāf-at*
mother-1SG both washed-3SG.F DEF-dishes and cleaned-3SG.F
l-bīst.
def-house
‘My mother [both] washed the dishes and cleaned the house.’

Optional, occurring in variation alongside various Arabic-derived counterparts such as *hatta ‘even’ or lākin ‘but’, is the contrastive correlative *balke ‘but [... also]’:

(18) *huwwa mā bass bāhūs balke šujjā’ ham.*
he neg only clever but brave too
‘He is not only clever but also brave.’

Further Persian borrowings that are generalized in Kh. Arabic are the concessive subordinating conjunctions *agarče and bā inke*, both ‘although/even though’, and the factual complementizer *ke* ‘that’:
(19) huwwa rāh lwaḥda l-ʾal-pārk agarēe umm-a
gall-at l-ʾa lā-yrāh.
said-3SG.F to-3SG.M NEG-go.3SG.M
he went.3SG.M alone to-DEF-park although mother-3SG.M
‘He went to the park alone, even though his mother told him not to go.’

(20) rayyāl-na bə-l-yūm xetab, bā įnke ṭalāṭašš sana
man-1PL in-DEF-day proposed.3SG.M although thirteen year
šumr-i sawwām rāhnamāī ubū-y qobal b-i
age-1SG third secondary school, father-1SG accepted for-3SG.M
‘When my husband proposed, although I was [just] thirteen years old, third year of secondary school, my father agreed.’

(21) tdr-in ke rayl-ʾač šala kəl-šī čađdaβ.
know-2SG.F COMP husband-2SG.F on everything lied.3SG.M
‘You know that your husband lied about everything.’

The latter, the Persian complementizer and relativizer ke, does not appear in non-factual (subjunctive) complements, where instead we find the Arabic (historical) relativizer l-šād or illi, which also continues to cover the function of a relativizer. Nonetheless, occasionally Persian ke is also found in the position of the relativizer:

(22) s̀bon uxā yāži ke huwwa w mart-a ḥnā ...
son brother Ghazi REL he and wife-3SG.M here
‘Ghazi’s nephew, who is here with his wife . . .’

From this we might assume a gradual process of convergence in steps, as follows: in stage 1, the Persian model of having an identical marker for complement clauses and relative clauses (ke) is copied into Kh. Arabic, with the effect of generalizing the relativizer l-šādilli (at the expense of the historical Arabic complementizer ‘imn-’) to cover the function of complementizer. The result is a convergence of patterns among the two languages. In stage 2, the actual Persian marker ke is adopted into Kh. Arabic in factual complement clauses, as seen in (21). The result is a split within Kh. Arabic between factual and non-factual complements, whereas the same marker is used in both languages to introduce factual complements. Finally, in stage 3, the beginnings of which are attested in the contemporary language, Persian ke infiltrates Kh. Arabic relative clauses as well, as seen in (22).
7. Constituent order

One change in constituent order has already been mentioned above, in Section 2: it concerns the shift in 'complex' attributive constructions, away from the Arabic norm, which allows an adjectival modifier to be separated from its head (by a nominal modifier of the complex noun phrase), toward the Persian-type constituent order, whereby each attribute must immediately follow its head. We repeat example (7) here:

(7) a. Khuzistani Arabic
   \textit{walad ĕ-čibir l-modir}
   boy DEF-big DEF-director
   'the director's big/eldest son.'

b. Standard Arabic (and other dialects)
   \textit{walad l-mudir l-ka\textsuperscript{b}ir}
   boy DEF-director DEF-big
   'the director's big/eldest son' (also: 'the big director's son')

c. Persian:
   \textit{pesar-e bozorg-e modir}
   boy-EZ big-EZ director
   'the director's big/eldest son'

A further issue related to the order of constituents in Kh. Arabic concerns the position of the copula-auxiliary \textit{lçânl}, which, in the composite past tense (pluperfect), follows the lexical verb: \textit{mâ-râyâh cân} 'he had not gone away' (Persian: \textit{na-râfte bi\textsuperscript{d}}) (see examples 10 and 11).

Noteworthy is also the flexible position of the causal conjunction \textit{čës} 'because'. Like its Persian counterpart \textit{čon}, it can also occupy the final position in the adverbial clause expressing cause:

(23)  a. Khuzistani Arabic
   \textit{lîes mâ-reh-ti  l-\textsuperscript{2}l-madrasa?}
   why NEG-went-2SG.F to-DEF-school
   cân \textit{fêd-dî xu\textsuperscript{t}âr čës.}
   was.3SG with-1SG guests because
   'Why didn't you go to school?'
   'Because I had guests.'
b. Persian
cerā be madrese na-raft-i?
why to school  NEG-went-2SG
mehmūn dāšt-am ēn.
guest  had-1sg because
‘Why didn’t you go to school?’
‘Because I had guests.’

Finally, we must consider what appears to be the beginning of a shift in word order, extending the contexts in which Object–Verb order is favoured to comply more frequently with the Persian type. Object–Verb order in Arabic is generally highly marked and is employed as a means to topicalize the direct object. Kh. Arabic makes use of such strategies, which include – unlike Persian, where OV prevails – the pronominal resumption of the object in a position following the lexical verb. Nevertheless, such constructions in Kh. Arabic do not necessarily express the topicalization of the object:

(24) la-bnayya d-dār naddaf-at-ha.
def-little.girl def-room cleaned-3SG.F-3SG.F
‘The little girl cleaned [it] the room.’

(25) hadān xālāt-i lisāns-han kazz-ann-a.
these aunts-1sg degree-3PL.F gained-3PL.F-3SG.M
‘My aunts received [it] their degree.’

8. Lexicon

The presence of numerous Persian lexical borrowings is a distinguishing feature of Kh. Arabic, setting it apart from other neighboring dialects of Arabic. Nevertheless, there is considerable sociolinguistic stratification in the use of Persian vocabulary among different groups of speakers (cf. Shabibi 1998). As the principal language of the public sphere, Persian supplies numerous lexical items in the domains of trade, institutions, tools, and other aspects of public and technical life (e.g. xarād-o-furās ‘trade’, pič guši ‘screwdriver’, lebās šū’i ‘washing machine’, etc.). In everyday vocabulary, Persian idioms are commonly calqued in Kh. Arabic, facilitated by the fact that those idioms themselves are often based on Arabic loan vocabulary in Persian, and so even more easily replicable in Kh. Arabic: Consider Kh. Arabic wāyad mamnūn, lit. ‘very grate-
ful’, in the sense of ‘thank you very much’, based on Persian xeyli māmnūn, or Kh. Arabic yarreti zahma, lit. ‘you have taken trouble [on my behalf]’, also an expression of gratitude, from Persian zahmat kešidi. Here, the fact that the languages already share a large part of their vocabulary (as a result of earlier, historical influence of Arabic on Persian), makes replication of lexical Matter redundant, and promotes in turn replication of idiomatic Patterns surrounding a pivotal word in the idiom that is already shared by both languages.

9. Conclusion

Matter replication of Persian material is found in Kh. Arabic primarily in the domain of lexical vocabulary, and in part in grammatical vocabulary, covering discourse markers that operate strictly on the interaction level (i.e. not conjunctions), focus particles, a correlative particle, a complementizer and relative particle, and concessive subordinating conjunctions. Pattern replication is most notable in the emerging change of constraints on word order (extension of marked word-order patterns), the favouring of analytic constructions and emergence of a new analytic past tense (pluperfect), and the reduction of overt marking of definiteness. Perhaps the most remarkable contact-induced change, one which strongly affects the typology of attribution in the language, is the identification of Kh. Arabic grammatical morphemes in attributive constructions – the Construct State marker (visible in the feminine singular only) and the definite article that appears between head and attribute – with the Persian attributive particle, and the consequent merger of two historically distinct attributive constructions – adjectival and nominal – into a single type, replicating the state of affairs in Persian.

Abbreviations

| ACC  | accusative               |
| COMP | complementizer           |
| CONS | construct state marker   |
| COP  | copula                   |
| DEF  | definite article         |
| DM   | discourse marker         |
| EZ   | Persian Ezāfe attributive marker |
| F    | feminine                 |
| INDEF| indefinite article       |
| M    | masculine                |
| NEG  | negation                 |
| PAST | simple past (perfective) |
| PL   | plural                   |
| POSS | possessive expression    |
| REL  | relative particle        |
| SG   | singular                 |
References

Ingham, Bruce

Shabibi, Maryam
Grammatical Borrowing in Cross-Linguistic Perspective

Edited by
Yaron Matras
Jeanette Sakel

Mouton de Gruyter
Berlin · New York