Arabic influence on Kurdish

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 change 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 West 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, around lake Urmiye 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 very hard to reliably determine since none of the sovereign countries make the relevant census information available. Table 1 this provides some cautious estimates based on various sources (especially Sirkeci 2005; Zeyneloğlu, Sirkeci & Civelek 2016; Ethnologue).[[1]](#footnote-1),[[2]](#footnote-2)

Table 1: Estimates of Kurdish population numbers

|  |  |
| --- | --- |
| Country | Population size |
| Turkey  Iraq  Iran  Syria | ca. 15.000.000  ca. 8.000.000  ca. 6.000.000  ca. 1.500.000 |

1. 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, Minorsky & MacKenzie 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 the historical Iraq and Syria, as well as in the Jazira area. 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 mixity 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, Minorsky & MacKenzie 2012: 449, 452, 456; James 2007: 115–120).

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 must have been 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, however 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.

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 the 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 some influence from Kurdish on the local Arabic dialects (cf. (Jastrow 2011) and see below §‎3.1).

As a result of these differing degrees and modalities of contact with Arabic, the influence of Arabic should be considered on at least two levels and separately for different country contexts where Kurdish is spoken. Of the two levels, there should be assumed a deeper contact influence shared in larger portions of Kurdish-speaking areas dating before the 20th century; and a more shallow level that is the result of the more recent societal bilingualism in Iraq and Syria. Likewise, while in Syria and Iraq the Arabic influence on Kurdish continues, this influence is largely replaced by the 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 level 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 restrictively – in Iraq, concerns also grammatical constructions and at least some of that contact influence could be due to imposition under SL-agentivity.

1. Contact-induced changes in Kurdish
   1. Phonology

The consonant inventory of Kurmanji is given in Table 2.[[3]](#footnote-3)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | bilabial | labio-dental | dental | alveolar | palatal | velar | uvular | pharyngeal | glotttal |
| Plos. | p’ p b |  | t’ t d | |  | k’k g | q |  | ’ |
| Fric. |  | f v | ş j | |  | x ẍ |  | ḥ ʿ | h |
| Affr. |  |  | ç’ ç c | |  |  |  |  |  |
| Nas. | m |  | n | |  | (ŋ) |  |  |  |
| Trill |  |  | r̄ | |  |  |  |  |  |
| Flap |  |  | r | |  |  |  |  |  |
| Approx. | w |  |  | | j |  |  |  |  |
| Lateral |  |  | l (dial. ł) | |  |  |  |  |  |

Table 2: Consonant phonemes in Kurmanji

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, (ii) it has velar nasal and velarized lateral phonemes.

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 source Semitic language for these sounds cannot be determined since Kurdish (or rather its ancestor languages) must have been in close contact with various Semitic languages for more than two millenia (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 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]](#footnote-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]](#footnote-5)

|  |  |  |  |
| --- | --- | --- | --- |
| Arabic | Northern Kurdish | Central Kurdish | Gloss |
| *ʕarab*  *maʕlūm*  *ʕadāla(t)*  *ṭābiʕ*  *maḥall*  *maḥšar*  *ḥākim*  *ḥammām*  *baḥr* | *ʿereb*  *meʿlûm*  *ʿedaḷet*  *tabiʿ/ṭabiʿ*  *miḥele*  *meḥşer*  *ḥakim*  *ḥemam*  *beḥr* | *ʿereb*  *meʿlûm[[6]](#footnote-6)*  *ʿedaḷet*  *ṭabiʿ*  *meḥal*  *meḥşer*  *ḥakim*  *ḥemam*  *beḥr* | ‘Arab’  ‘evident’  ‘dependent’  ‘neighborhood’  ‘neighborhood’  ‘resurrection (day)’  ‘judge, governor’  ‘bath’  ‘sea’ |

Table 3: Loanwords with retained pharyngeals in Kurdish

Some loanwords with original pharyngeals are reanalysed into 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 pharyngealof Ar *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 both Arabic and non-Semitic loanwords. Thus in most of Northern Kurdish the Arabic word *ʔarḍ* ‘earth’ appears with a 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 been expanded also into inherited native Iranian lexicon, especially in Northern Kurdish. However, unlike in Arabic loanwords, the fluctuation between pharyngeal and non-pharyngeal uses of the words among the dialects (sometimes in immediate geographic proximity) is much more 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.

|  |  |  |  |
| --- | --- | --- | --- |
| Persian | Northern Kurdish | Central Kurdish | Gloss |
| *abr* | *ʿewr* | *hewr* | ‘cloud’ |
| *zabān* | *ʿezman ~ ziman* | *ziman* | ‘language’ |
| *āsemān* | *ʿesman* | *asman ~ hasman* | ‘sky’ |
| *xošk* | *ḥişk ~ hişk* | *wişk* | ‘dry, hard’ |
| *haft* | *ḥeft ~ heft* | *ḥewt* | ‘seven’ |
| *hašt* | *ḥeşt ~ heşt* | *heşt* | ‘eight’ |
| *bahašt* | *biḥeşt* | *beheşt* | ‘paradise’ |

Table 4: Pharyngeal sounds in native Iranian lexical items

More striking, however, is the emergence of a voiced pharyngeal in a subset of 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 Northern Kurdish 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’).[[7]](#footnote-7)

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:

|  |  |  |
| --- | --- | --- |
| Arabic | Northern/Central Kurdish | Gloss |
| *ṭaʕm*  *ṭāʔir*  *baṭṭāl*  *ð̣ulm*  *ḍābit*  *ṣūfī*  *ṣāfī* | *ṭam* ~ *ṭeʿm / tam*  *ṭeyr*  *beṭal*  *ẓulm*  *ẓabit*  *ṣofî*  *ṣaf/ṣafî* | ‘taste’  ‘bird’  ‘empty, cancel’  ‘oppression’  ‘clerk’  ‘devotee, Sufi’  ‘clear’ |

Table 5: Arabic loanwords with emphatic consonants in Kurdish

In the deeper-layer loanwords the Arabic interdental and alveolar emphatics are merged into the alveolar phoneme *ẓ* in Kurdish, although a less emphatic one. But in the present-day Iraqi and Syrian Kurdish speech, especially the 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 without their original emphatic consonants, reanalysed thus as corresponding plain phonemes – similar to Persian – as in the items in Table 6.

|  |  |  |  |
| --- | --- | --- | --- |
| Arabic | Northern Kurdish | Central Kurdish | Gloss |
| *xāṭir*  *ṭaraf*  *šayṭān*  *ḍaʕīf*  *ḥāḍir*  *qaṣṣāb*  *fasīḥ*  *ṣabr* | *xatir*  *teref*  *şeytan*  *zeʿîf*  *ḥazir*  *qesab*  *fesîḥ*  *sebr~ṣebr* | *xatir*  *teref*  *şeytfn*  *zeʿîf*  *ḥazir*  *qesab*  *fesîḥ*  *sebr* | ‘mind’  ‘side, direction’  ‘devil’  ‘weak’  ‘ready’  ‘butcher’  ‘clear’  ‘patience’ |

Table 6: Arabic loanwords with lost emphatics in Kurdish

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, pointing to a higher degree of integration into the phonological system of the language.

|  |  |  |
| --- | --- | --- |
| 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’ |
| *ḍiṙî* | *-* | ‘blackberry bush’ |

Table 7: Emphatic consonants in native Iranian lexical items

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 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 southern 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 consonant 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 Kurdish. 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), Barry (forthcoming) and Anonby (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 *h* sound in 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, although acknowledging the role of language contact with Arabic in the initial stage, it postulates a phonetic mechanism of language-internal development of pharyngealization that captures the expansion of pharyngeals onto historically non-pharyngeal inherited and loan vocabulary 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* and *a* and *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: (1) *civat* ‘community, company’,(2) *cimat* ‘the assembly of prayers in a funeral’,(3) *cemaʕet* ‘community’. The first form is probably the result of an early borrowing right after the centuries of islamicization of the Kurds, as the fricativization of the bilabial nasal was active then (as seen also in *silav* ‘greeting’ from Arabic *salām*, cf. 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 sensitive phonetic environment (i.e. 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 of 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. Anonby forthcoming; Haig 2007), since, as noted by Haig (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). The phonemes and example words with their sources are given below:

bilabial stop /p/ *parčāye* ‘piece’ < Tr. *parça*

voiced labio-dental fricative /v/ *davare* ‘ramp’ < Kr. *dever* (f) ‘place’

voiceless affricate /č/ *čǝqmāq* ‘lighter’ < Tr. *çakmak*[[8]](#footnote-8)

voiced palatal fricative /ž/[[9]](#footnote-9) *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/[[10]](#footnote-10) *xōrt* ‘young man’ < Kr. *xort*

These additions into the phoneme inventory of the Anatolian Arabic are clearly 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.

* 1. 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 prefix (in turn: *mī-, di-, gǝ-*, *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 be diffused from Kurdish and Persian into the other languages of the zone, including Arabic (which in its standard grammar does not have such forms, cf. Ryding (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 mainly 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 propensity in the languages of Anatolia, Haig (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 – previously not 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 into 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 2016: 5). Thus 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 (Badini) of Iraq uses masculine gender for integrating modern vocabulary items into the language. The following modern lexical borrowings (underlined) are all assigned to masculine gender in Badini Kurmanji of Iraq. Note that the gender of the nouns is visible in the *ezafe* (see §‎3.3) and oblique suffixes.

1. Badini dialect of Kurmanji in Iraq (from Badini media outlets)

sîstem-ê endroyd-ê

system-ez.m android-obl.f

‘Android system’



serok-ê parleman-î

president-ez.m parliament-obl.m

‘the president of the parliament’

form-ê têgehiştin-ê

form-ez.m understanding-obl.f

‘the form of understanding’

moral-ê diyalog

moral-ez.m dialogue

‘the moral of dialogue’

proj(e)-ê av-ê

project-ez.m water-obl.f

‘the water project’

prensîp-ê hevwelatîbûn-ê

principle-ez.m citizenship-obl.m

‘the principle of citizenship’

Now all of these lexical borrowings exist also in the Kurmanji spoken in Turkey, but they are systematically used with feminine gender. For instance the phrase in ‎(1)‎b 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, most of such modern lexical borrowings in Arabic are assigned to masculine gender. The masculine gender assignment in Badini 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 Badini. However, this contact influence must have been established relatively recently, since earlier technical borrowings such as *têlevizyon* and *radyo* are treated as feminine nouns, despite being masculine in Arabic.

* 1. 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 order of a possessor in a multiple-modifier noun phrase in Northern Kurdish.

As is well known, a number of Western Iranian languages (Middle Persian, Persian, Tajik, Luri, Kurdish, Zazaki, etc.) employ a vocalic bound morpheme for linking post-head modifiers in a noun phrase, called *ezafe* or *izafe* (from Arabic *iḍāfa* ‘joining, addition’), as in ‎(2) and ‎(3).

1. Persian (personal knowledge)

xāna-e bozorg

house-ez big

‘(the) big house’

1. Northern Kurdish

xanî-yê mezin

house-ez.m big

‘the big house’

The *ezafe* 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/pronominal forms used in chain *ezafe* constructions with multiple modifiers (and some other predicative functions, cf. Haig (2011); Haig and Öpengin (2018)). In most of the West Iranian languages, noun phrases with multiple modifiers have their head noun first followed by qualitative modifiers then by possessive ones, as in ‎(4) and ‎(5). This is also the order in Middle Persian, as in ‎(6), where Tsabolov (1994: 122) thinks such constructions may be regarded as prototypes of the *ezafe* constructions of modern West Iranian languages.

1. Persian (personal knowledge)

xāna-e bozorg-e Malek

house-ez big-ez pn

‘Malek’s big house’

1. Central Kurdish (Tsabolov 1994: 122)

kiç-î cwan-î hewal-î min

daughter-ez beautiful-ez friend-ez 1sg.poss

‘my friend’s beautiful daughter’

1. 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, whereby 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.

1. 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).

1. Arabic (Tsabolov 1994: 123)

miḥfað̣atu ṭ-tālibi l-jadīdatu

bag.nom def-student.gen def-new.f.nom

‘the student’s new bag’

1. 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 *ezafes* 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 secondary *ezafe* are different (with masculine *-î* and feminine *-e*). In Tsabolov’s view, the century-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 *ezafe* through the “weakening” of the primary ones (morphological), because, he thinks, such distinct forms “were necessary for correlating each attribute in an [*ezafe*] 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, *ezafe* forms in Northern Kurdish distinguish gender/number that already correlate the modifiers with their head nouns, and on the other hand, in the majority of Northern Kurdish dialects the primary and secondary *ezafes* are formally identical. The change in form is an instance of vowel raising (a > e, ê > î) that is observed also 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 a 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 *ezafe* 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 *ezafe* 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 *ezafe* 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. Remember that in Central Kurdish the possessor in a chain *ezafe* construction is positioned at the end of the noun phrase, as illustrated in ‎(5). However, here the proper name occurs right after the head noun and before the qualitative modifier, as in ‎(10) and ‎(11).

1. Central Kurdish (official signage)

qutabxane-y Qemeryan-î seretayî

school-ez pn-ez primary

‘Qamaryan primary school’

1. Central Kurdish (official signage)

firokexane-y Hewlêr-î nawneteweyî

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 sytax of chain *ezafe* 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 as in ‎(12).

1. 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.

* 1. 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 (2004: 97), the process of lexical convergence initially took place in Persian between the ninth and thirteenth centuries when an important 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 geographically closer area 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 on the 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 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):[[11]](#footnote-11)

*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*

*fikr* ‘thought, idea’ < Ar *fikr*

*fêkî/fêqî* ‘fruit’ < Ar *fākiḥa*

*ḥal* ‘condition’ < Ar *ḥāl*

*ḥazir* ‘ready’ < Ar *ḥāḍir*

*şol/şuẋul* ‘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 (Öpengin & Haig 2014), spoken around Marin 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 Arabic *ṭafa* or standard *ṭafiʔa*), *şiteẋlîn* ‘speak-intr.inf’ (from dialectal Arabic *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, 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 *ʕammah*; this item does not exist in Sorani), *xalet/xaltî* ‘maternal aunt’ (Ar *ḫālah*), *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*), *ḥubr* ‘ink’ (Ar *ḥibr*), *ḥemam* ‘bath’ (Ar *ḥammām*), *ḥewş* ‘yard’ (Ar *ḥawš)*, *mexmer* ‘velvet’ (Ar *muḫmal*). Some occupational terms from Arabic are *neqş* ‘embroidery’ (Ar *naqš* ‘painting, drawing’), *ḥedad* ‘blacksmit’ (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 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*), *meḥkeme* ‘court’ (Ar *maḥkama*), *ḥukûmet* ‘government’ (Ar *ḥukūma*), *mekteb* ‘school’ (Ar *maktab* ‘office, desk’), *deʿwā* ‘request, court case’ (Ar *daʕwa* ‘request, invitation’ and *daʕwā* ‘court case’), *qanûn* ‘law’ (Ar *qānūn*).

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 *waʿð̣*), *ḥelal* ‘permitted’ (Ar *ḥalāl*), *ḥeram* ‘forbidden’ (Ar *ḥarām*), *ruḥ* ‘soul, spirit’ (Ar *rūh*), *tizbî* (Sorani *tezbêḥ*) ‘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 *asīr*), *dinyā* ‘world’ (Ar *dunyā*), *ḥesab* ‘count, calculation’ (Ar *ḥisāb*), *ḥîle* ‘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 *daʕwa* ‘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 Christians 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’ + *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 the verb forms. This takes place either through morphological integration or syntactic composition. In morphological integration, the Arabic root (whether substantive or verbal) is taken as the stem onto which 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’).[[12]](#footnote-12) The verb *qelandin* ‘to roast; to uproot’ has two sources as Ar *qalā* and *qalaʕa,* respectively, which explains its polysemy in Kurdish.

In syntactic composition, on the other hand, a compound verb[[13]](#footnote-13) 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 the following:

*qedr* ‘respect’ (Ar *qadr*)+ *girtin* ‘to hold’ = ‘to respect’

*silav* ‘greeting’ (Ar *salām* ‘peace’) + *dan* ‘to give’ = ‘to greet’

*teʿn / ṭan* ‘scolding’ (Ar *ṭaʕn* ‘piercing’)+ *dan* = ‘to criticize’

*qedeẍe* ‘forbidden’ (Ar *qadaḥa* ‘to rebuke’) +*kirin* ‘to do’ = ‘to forbid’

*qesd* ‘intention’ (Ar *qasd*) + *kirin* = ‘to head for’

*zeʿîf* ‘weak’ (Ar *ḍaʕīf*) + *bûn* ‘to be’ = ‘to become slim’

The principle behind 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 either of the 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 Arabic *ʕimil* ‘to do’), *şuẋulîn* (Ar *šuɣl* ‘work’), *fikirîn* (Ar *fikr* ‘thought’) are seen in the southeastern dialect (also called Badini) 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 the 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). Following are 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 province 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 ‎(13).

1. Southern dialect of Northern Kurdish (Media)[[14]](#footnote-14)

nema di-kar-im veger-im welêt

no.longer ind-be.able.prs-1sg return.prs.sbjv-1sg country

'I no longer can return to the homeland'

There is an immediately corresponding adverbial form *mā ʕād* ‘no longer’ in Arabic, which is based on 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 ‎(14).

1. Central Kurdish (Media)[[15]](#footnote-15)

polîs hellsa be kokirdinewe-y zanyarî

police rise.pst.3sg with collecting-ez information

‘The police undertook (the task of) collecting information.’

1. Northern Kurdish (Media)[[16]](#footnote-16)

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 has also its 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.

1. Conclusion

Contact with Arabic, which started in the early medieval period (ca. 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 quite 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 north of Raqqa). 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. It has some repercussions on the expansion of the phonological inventory of the language and it is shared across most of Kurdish. 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 and 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 could also be related to bilingualism and language configuration in historical perspective. That is, the absence of imposition (as 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 such imposition is observed in the Kurmanji of the Jazira region, which is known to have had greatest speaker contact between Kurdish and Arabic speech communities. To the contrary, the widespread bilingualism and Arabic-dominant linguistic configuration in Syria and Iraq for at least one century has likely 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 take place.

Given the limitations of a first attempt, much is yet to be explored regarding Kurdish–Arabic language contact. Particularly, 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 very 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 the shortcomings and errors.

Abbreviations

Ar: Arabic

BCE: before common era

ca.: circa

def: definite

drct: directional

dial.: dialectal

ez: *ezafe*

f: feminine

gen: genitive

Kr.: Kurdish

ind: indicative

indf: indefinite

inf: infinitive

intr: intransitive

ipfv: imperfective

m: masculine

neg: negative

nom: nominative

pl/pl.: plural

pn: proper noun

poss: possessive

prs: present

pst: past

pvb: preverb

sbjv: subjunctive

sec: secondary or pronominal *ezafe*/linking element

sg/sg.: singular

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 Haig & Erik Anonby (eds.), *Current Topics in Kurdish Linguistics* (Bamberg Studies in Kurdish Linguistics 1). 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. (Ed.) P. Bearman, Th. Bianquis, E. Bosworth, E van Donzel & W.P. Heinrichs. *Encyclopaedia of Islam, Second Edition*. Leiden: Brill.

Chyet, Michael. 2003. *Kurdish–English dictionary*. New Haven: Yale University Press.

Coetsem, Frans van. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

Coetsem, Frans van. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.

Coghill, Eleanor. 2016. *The rise and fall of ergativity in Aramaic: cycles of alignment change* (Oxford Linguistics 21). First edition. Oxford: Oxford University Press.

Edwards, Bela B. 1851. Note on the Kurdish language. *Journal of the American Oriental Society* 2. 120–123.

Folmer, M L. 1995. *The Aramaic Language in the Achaemenid Period. A Linguistic Study in Variation*. Leuven: Peeters.

Haig, Geoffrey. 2002. Complex predicates in Kurdish: Argument sharing, incorporation, or what? *Sprachtypologie und Universalienforschung / Language typology and universals* 55 (1). 15–48.

Haig, Geoffrey. 2007. Grammatical borrowing in Kurdish (Northern Group). In Y Matras & Janet Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 165–184. Berlin: Mouton de Gruyter.

Haig, Geoffrey. 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 Volume 1: Sino-Tibetan and Iranian Languages*. Amsterdam: John Benjamins.

Haig, Geoffrey. 2014. East Anatolia as a linguistic area? Empirical and conceptual issues. In Lale Behzadi, Patrick Franke, Geoffrey Haig, Christoph Herzog & Birgitt Hoffmann (eds.), *Bamberger Orient Studien*, vol. I, 13–36. University of Bamberg Press. Bamberg.

Haig, Geoffrey L.J. & Ergin Öpengin. 2015. Gender in Kurdish: Structural and socio-cultural dimensions. In Marlis Hellinger & Heiko Motschenbacher (eds.), *IMPACT: Studies in Language and Society*, vol. 36, 247–276. Amsterdam: John Benjamins Publishing Company. doi:10.1075/impact.36.10hai. https://benjamins.com/catalog/impact.36.10hai (18 June, 2019).

Haig, Geoffrey & Ergin Öpengin. 2014. Kurdish: A critical research overview. *Kurdish Studies* 2(2). 99–122.

Haig, Geoffrey & Ergin Öpengin. 2018. Kurmanji Kurdish in Turkey: structure, varieties and status. In Christiane Bulut (ed.), *Linguistic Minorities in Turkey and Turkic-speaking Minorities of the Peripheries*, 157–230. Wiesbaden: Harrassowitz.

Ibrahim, Iman. 2016. Gender Assignment to Lexical Borrowings by Heritage Speakers of Arabic. *Western Papers in Linguistics / Cahiers linguistiques de Western* 1(2). https://ir.lib.uwo.ca/wpl\_clw/vol1/iss2/1.

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* (117–118). 101–126. doi:10.4000/remmm.3331.

Jastrow, Otto. 2011. Turkish and Kurdish influences in the Arabic Dialects of Anatolia. *Türk Dilleri Araştırmaları* 21(1). 83–94.

Kahn, M. 1976. *Borrowing and regional variation in a phonological description of Kurdish*. Ann Arbor, Michigan: Phonetics Laboratory of the University of Michigan.

Kent, Roland. 1953. *Old Persian*. New Haven, Connecticut: American Oriental Society [second, revised edition].

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. I. Oxford: Oxford University Press.

Matras et al., Yaron. 2016. *The Dialects of Kurdish*. Manchester: University of Manchester. http://kurdish.humanities.manchester.ac.uk/ (2 July, 2018).

Matras, Yaron. 2010. Contact, Convergence, and Typology. In Raymond Hickey (ed.), *The Handbook of Language Contact*, 66–85. Oxford: Wiley-Blackwell.

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. doi:10.1515/ijsl-2012-0053.

Öpengin, Ergin. 2016. *The Mukri variety of Central Kurdish: grammar, texts, and lexicon* (Beiträge Zur Iranistik Band 40). Wiesbaden: Ludwig Reichert Verlag.

Öpengin, Ergin & Geoffrey 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. *Encyclopædia Iranica*. New York: Encyclopædia Iranica Foundation. http://www.iranicaonline.org/articles/kurdish-language-i.

Perry, John R. 2004. 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 Diffusion: Case Studies from Iranian, Semitic and Turkic*. New York: 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 and Anthropology.

Ripper, Thomas. 2012. *Diyarbekir Merwanileri: İslami Ortaçağ’da Bir Kürt Hanedanı*. (Trans.) Bahar Şahin Fırat. İstanbul: Avesta.

Ryding, Karin C. 2014. *Arabic: a linguistic introduction*. Cambridge, United Kingdom: Cambridge University Press.

Sirkeci, Ibrahim. 2005. War in Iraq: Environment of Insecurity and International Migration. *International Migration* 43(4). 197–214. doi:10.1111/j.1468-2435.2005.00338.x.

Teyran, Feqiyê. 2011. *Dîwana Feqiyê Teyran*. (Ed.) Seîd Dêreşî. 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 Diffusion: Case Studies from Iranian, Semitic and Turkic*, 65–77. New York: Routledge.

Zeyneloğlu, Sinan, Ibrahim Sirkeci & Yaprak Civelek. 2016. Language shift among Kurds in Turkey: A spatial and demographic analysis. *Kurdish Studies* 4(1). 25–50. doi:10.33182/ks.v4i1.405.

1. Main URL: https://www.ethnologue.com/language/kur (Last accessed May 20th, 2018) [↑](#footnote-ref-1)
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, Sirkeci & Civelek 2016). [↑](#footnote-ref-2)
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 (last accessed May 28th, 2018) [↑](#footnote-ref-3)
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 et al. 2016)). The Sorani lexical items are from Öpengin (2016) and the popular press. [↑](#footnote-ref-4)
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. [↑](#footnote-ref-5)
6. Note that vowel length is marked with a circumflex in the Kurdish Latin alphabet. [↑](#footnote-ref-6)
7. URL: http://kurdish.humanities.manchester.ac.uk/pharyngeal-retentionloss-animal/ (last accessed: May 25, 2018) [↑](#footnote-ref-7)
8. 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/closing in the vowel of the first syllable (cf. also *qeymaẍ* ‘cream’ from Turkish *kaymak*) are typical of Kurmanji spoken in the region. [↑](#footnote-ref-8)
9. Note that the reflex of Arabic ج in this variety is /ǧ/, not /ž/. [↑](#footnote-ref-9)
10. Note also that the original Arabic diphthongs \*ay and \*aw are preserved in this variety, not monophthongized to /ē/ and /ō/. [↑](#footnote-ref-10)
11. The most important source for the lexical items in this section, together with the information of 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 the shortcomings. [↑](#footnote-ref-11)
12. Kurdish possesses a number of preverbs such as *ve-, 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. [↑](#footnote-ref-12)
13. Here the term *compound verb* is employed in a pre-theoretical sense, regardless of whether or not the given complex verb is considere to form a compound. See Haig (2002) for a discussion of complex verb forms in Kurdish. [↑](#footnote-ref-13)
14. 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/ (Last accessed March 28th, 2019) [↑](#footnote-ref-14)
15. Journalistic piece: http://www.kurdistan24.net/so/news/5ca67132-7a7f-4840-bfb4-dea5bf25ea2e (Last accessed March 28th, 2019) [↑](#footnote-ref-15)
16. Journalistic piece: http://hncand.blogspot.com/2015/04/celadet-bedirxan-bi-tene-sere-xwe-bi.html (Last accessed March 28th, 2019) [↑](#footnote-ref-16)