

To appear in: *Encyclopedia of Arabic Language and Linguistics*. K. Versteegh, editor. Leiden: Brill.

Gypsies, Arabic of

Yaron Matras, *University of Manchester*

‘Gypsies’ is an ambiguous term. It is used on the one hand as a universal term to denote ethnically and linguistically diverse populations of commercial nomads (also known as service nomads, itinerants or peripatetics). In a more restricted sense, it often refers specifically to the *Řom* or *Romanies* of Europe, a population of Indian origin whose language is (or, in the case of some communities, was) a dialect of Romani – *řomani čhib* or *řomanes*, as it is usually referred to by its speakers (cf. Matras 2002). A further, mixed reading of the term ‘Gypsies’ might include populations of commercial nomads outside of India who, like the Romanies, are of Indian origin, but who speak an Indian language that is not a dialect of Romani. This includes the populations known as *Dom* (also *Duman*, *Qurbāt*, *Karači*) in the Middle East, whose language is known as *Domari* (cf. Matras 1999), populations like the *Jat* of Afghanistan (Rao 1995) or the *Ḍum* of the Hunza valley (Lorimer 1939), who speak Central Indian languages, and perhaps also the *Lom* or *Boša* of Anatolia and Armenia (Finck 1907, Patkanoff 1908), who speak Armenian but retain a distinct in-group vocabulary of Indo-Aryan origin, known as *Lomavren*. We will follow the broader interpretation of the term for the purpose of this description, associating it with populations of commercial nomads, irrespective of origin or ethnicity, in the Arabic-speaking area.

Linguistically, there are three separate phenomena that potentially merit attention under the above heading: First, the use of an in-group special vocabulary, of a limited size and usually of limited communicative functions, by groups whose everyday family and community language is a form of Arabic. Second, the incorporation of Arabic structures into the speech of peripatetic communities that constitute linguistic minorities in the stricter sense, that is, who speak a language other than Arabic amongst themselves, but use Arabic in interaction with outsiders. Third, the kind of Arabic dialect, sociolect, or ethnolect used by mi-

nority peripatetic communities. In the absence of any data on the third phenomenon, we will limit our attention to the first two.

The use of special vocabulary to cover everyday, i.e. non-technical meanings (or ‘basic’ vocabulary) in group-internal communication is a well-known universal feature of peripatetic communities, and is documented among diverse communities in many regions and in different continents. Examples are English *Cant*, Hiberno-English *Gammon*, Spanish *Germanía*, German *Rotwelsch*, Czech *Hantýrka*, Dutch *Bargoens*, and more. Such speech varieties are often referred to as ‘secret languages’, and, to the extent that they draw on vocabulary deriving from a particular second language, as ‘mixed languages’. Their status as full-fledged languages, however, is disputable. Essentially we are dealing with a fixed, albeit often flexible and volatile set of lexical items covering a limited range of meanings, and so with something that might rather be defined as a ‘disguised vocabulary’ – a reservoir of lexical items that are known only to group members. Its primary function is to exclude outsiders from key portions of the discourse, by disguising key meanings in the sentence. Sometimes special vocabularies are also used to establish group membership, to flag group identity, or to mark out the dichotomy between insiders and outsiders (cf. Hanna 1993:80–83). Compared to ‘languages’ in the normal sense, special vocabularies are thus structurally and functionally restricted. Grammatical structures usually remain unaffected by the special vocabulary. The occasional confusion of special vocabularies with pidgins or creoles is therefore incorrect.

Only limited documentation exists on Arabic-based special vocabularies. It is nevertheless clear that different sets of vocabularies are used by different groups, though there is quite often some overlap. A clear-cut taxonomy relating groups to types of special lexicon is made difficult both by the paucity of material, and due to the fact that there is only partial overlap between group-name and the composition of the various special vocabulary sets. Thus, any two groups known respectively by names such as *Ġajar*, *Ḥalab*, *Nawar*, *Qurbāt*, or *Bahlawān* may have either identical, partly overlapping, or even entirely different special vocabularies.

The special vocabulary items themselves may be divided into different types. The first type are language-internal formations that have their origin in Arabic itself, and derive from a conscious, deliberate attempt to disguise everyday Arabic words. This procedure is well-attested in other special vocabularies (e.g. ‘Pig Latin’), and is sometimes referred to as ‘cryptolalic formation’. Vycichl (1959) had already presented an overview of different cryptolalic techniques in what he calls the ‘slang’ of the *Ḥalab is-Sūdān*, whom he encountered in the vicinity of Luxor, Egypt. They tend to match cryptolalic formations that appear in wordlists collected among other peripatetics, for example by Newbold (1856) among the *Ḥalab* of Egypt, by von Kremer (1860) among the *Ġajar* of Upper Egypt, by Hanna (1995) among the *Ġajar* of Cairo, or by Streck (1995) among the *Ḥalab* of Sudan. Although cryptolalic formations have their origin in lexical camouflage strategies, the fact that we encounter the same items in various locations and among different groups indicates that the formation strategies are not usually on-the-spot productive techniques. Rather, they belong to the diachrony of the word, having been formed at some earlier point and then transmitted from one generation of users to another.

Morphological distortion of words is a common cryptolalic formation. Vycichl (1959) mentions the pattern *fū‘āl – ṭurrāg* ‘road’ (*ṭarīg*), *ṭubbāx* ‘cooked vegetables’ (*ṭabīx*). A widespread pattern is the insertion of the root of the target word into a special derivation pattern involving *m-* and a suffix *-iš*: *mubwābiš* ‘door’ (*b.w.b*), *muftāhiš* ‘key’ (*f.t.ḥ*) (Vycichl 1959); *menāhriš* ‘day’ (*n.h.r*), *maḥrāriš* ‘hot’ (*ḥ.r.r*), *mebrādiš* ‘cold’ (*b.r.d*) (Newbold 1856); *maxšābeš* ‘wood’ (*x.š.b*), *midhābeš* ‘gold’ (*d.h.b*), *migbāliš* ‘mountain’ (*g.b.l*), *mutwariš* ‘bull’ (*t.w.r*), *minxališ* ‘palm’ (*n.x.l*) (von Kremer 1860); *maṣabī‘aš* ‘finger’ (*ṣ.b.ʿ*), *madahaībš* ‘gold’ (*d.h.b*), *maxtīamš* ‘ring’ (*x.t.m*) (Hanna 1993). There are corresponding feminine forms: *mubṭānše* ‘belly’ (*b.t.n*), *misnānše* ‘tooth’ (*s.n.n*), *muwdānše* ‘ear’ (*w.d.n*) (Vycichl 1959), *mubṣālše* ‘onion’ (*b.ṣ.l*), *mubgarše* ‘cow’ (*b.g.r*). *mudānše* ‘ear’ (*w.d.n*) (von Kremer 1860). The two camouflage morphemes may also appear independently. Hanna (1993) notes *manūra* ‘light’ (*n.w.r*), as well as *ma’’aṣbāḥ* ‘morning’ (*ṣ.b.ḥ*), *ma’’akbīr* ‘big’ (*k.b.r*), *ma’’aṣḡīr* ‘small’ (*ṣ.ḡ.r*). Plain addition of a camouflaging suffix *-ayiš* is noted by von

Kremer (1860): *ḥuṣānāyīš* ‘horse’, *šagarāyīš* ‘tree’, *ḥadīdāyīš* ‘iron’, *dībāyīš* ‘wolf’, cf. also *aswādīš* ‘black’ (Newbold 1856). While *m-* is clearly the Arabic nominal/participial marker, with *ma-* deriving from the exclamative/emotive form (‘what a ...’), the suffix *-iš*, which Vychičl (1959:224) speculates might be an Indo-European nominative ending, is strongly reminiscent of the Domari nominaliser *-iš*: cf. Domari *mang-* ‘to beg’, *mangiš* ‘begging’. Littman (1920) on the other hand derives it from *šī* ‘thing’. Phonological distortions are widespread with numerals: *tulit* ‘three’, *rūbi* ‘four’, *xūmis* ‘five’, *sutet* ‘six’, *sūbi* ‘seven’, *tūmin* ‘eight’, *tiwa* ‘nine’, *ušīr* ‘ten’ (von Kremer 1860, Streck 1995).

Another widespread cryptolalic strategy is the functionalisation of figurative and metaphorical constructions. Von Kremer (1860) notes *mumešayāt* ‘feet’ (from *m.š.y* ‘walk’), and paraphrases such as *ma’āḥli* ‘dates’ (‘sweet stuff’), *el-ma-’asfar* ‘gold’ (‘the yellow stuff’), and *magaswade* < *ma-’aswad* ‘coffee’ (‘black stuff’). Metaphors combined with camouflage morphology are found in *baḥarāyīš* ‘north’ (*b.ḥ.r.* ‘sea’ = toward the Mediterranean), *kiblayīš* ‘south’ (toward Mecca, the *kibla*). In Vychičl’s (1959) list, a special morphological derivation – *mukaf’al* – is employed with metaphorical associations: *mukabwad* ‘eggs, milk’ (*b.y.ḍ* ‘white’), *mukaswade* ‘coffee’ (*s.w.ḍ* ‘black’), *mukaḥmar* ‘one pound’ (*ḥ.m.r.* ‘red’ = ‘gold coin’), *muganwara* ‘lamp, light, fire’ (*n.w.r* ‘light, fire’), *mukabwar* ‘fish’ (*būri*, a Nile fish). Other metaphorical extensions include forms such as *yamūy* ‘to drink’ (*mūy* ‘water’), or *sabsab* ‘hair’ (*ysibsib* ‘to comb’) (Hanna 1993). Word-derivation may combine figurative or paraphrase formations with generic or dummy words such as *māx* ‘thing’ or *anta* ‘place’: *māx l-mōya* ‘well’ (‘water-thingy’), *anta l-kabīr* ‘town’ (‘big place’), *anta ṣ-ṣaḡīr* ‘village’ (‘small place’). Some vocabulary items appear to derive from local usages and idiomatic expressions. Winkler (1936:389; cited in Streck 1995:300) for example derives the *Ḥalab* word for money, *buṭūqa* or *baṭqa*, from the Cairene name of a Spanish coin – *abū ṭāqa*, which pictures fields, that appear as ‘windows’. Sudanese *Ḥalab* *kūšī* ‘black African’ (Vychičl 1959) can be derived from *Kush*.

Internal (Arabic-based) cryptolalic formations are rarely the sole basis of the special lexicon. Most vocabulary sets also show words that appear to be of for-

eign origin, though in many cases their etymology remains unclear. Widespread non-Arabic items of unknown origin in the speech of the *Ġajar* and *Halab* of Egypt and Sudan for instance include *raxwa* ‘food’ and *raxxa* ‘to eat’, *watab* ‘to come’ and *wattab* ‘to bring’, *kodde* ‘woman’, *anta* ‘place’, *dāzī* ‘policeman’, *xušni* (pl. *xašāna*) ‘non-Ġajar’, *hidīd* ‘night’, *māx* ‘thing’ (also ‘one’), and more. Recognizable etymologies show a range of different contributor languages. Vyčichl (1959) names Aramaic *damax* ‘to sleep’ and *muṭallim* ‘blind’, and Nubian *tōd* ‘boy’, *buru* ‘girl’, *amanga* ‘water’. The Nubian influence can be attributed to contacts with other Sudanese peripatetic groups that are or were Nubian-speaking. The Aramaic component on the other hand is found in special vocabularies of peripatetic groups as far as Iran and Afghanistan. It is likely to derive originally from the use of Aramaic as trade language or lingua franca in the region, though the concrete diffusion of individual Aramaic-origin lexemes into various special vocabularies of present-day peripatetic groups may be much more recent, and attributable simply to contact among the various groups and to vocabulary borrowings or admixture of the special lexicon sets. A small number of words of Iranian origin are likely to have been adopted in a similar fashion.

While items such as *piyaz* ‘onion’, *gošt* ‘meat’, or *deh* ‘ten’ could be of either Persian or Kurdish origin, others, such as Newbold’s (1856: 295) *Nawar* numerals *suso* ‘three’ and *čar* ‘four’ point somewhat more clearly to a Kurdish origin. The source of at least some of this vocabulary may in fact be a peripatetic group of Kurdish origin: The *krād* ‘Kurds’ of the Palestinian West Bank are itinerant metalworkers who speak Arabic, but have a special vocabulary that is based partly on Domari, and partly on Kurdish. Palestinian Domari itself also contains many Kurdish loans, which are indicative of a prolonged stay in Kurdish-speaking territory prior to immigration into the present location. Among the isolated items of Turkish origin we find *gemi* ‘ship’, and especially widespread *kapi* ‘door’, which also appears in Domari. A number of items in Newbold’s lists of *Nawar* and *Ġajar* words, notably *namak* ‘salt’ and *thoraki* ‘a little’, appear to be of Hindi origin, and may have similarly been transmitted into the special vocabularies of Egypt via other special lexicons.

An interesting contribution to the Arabic-based special lexicon is that made by (European) Romani to the vocabulary of *Ġajar* of Egypt, as documented first by Newbold (1856) and later confirmed by Streck (1995) for the *Ġajar* of Sudan. The two vocabularies share many similarities, including the same deviations from the common Romani shape of the word, for example *gaziye* ‘wife’ (Newbold) *qazihe* ‘woman’ (Streck), Romani *gaži* ‘woman, wife’; *marey* (Newbold) *mari*’ (Streck) ‘bread’, Romani *maro*; *reibo* ‘king’ (Newbold) *raibó* ‘policeman, non-Gypsy’ (Streck), Romani *raj* ‘non-Gypsy official’. Both vocabularies are mixed and contain also non-Romani items, including, in Newbold’s list, widespread items like *kuddi* ‘mother’ (elsewhere ‘woman’), as well as items derived from Domari, such as *bakra* ‘sheep’ (Domari *bakra*, Romani *bakro*), *sir* ‘head’ (Domari *sir*, Romani *šero*), *kustúr* ‘hand’ (Domari *xastúr* ‘your hand’, Romani *vast*). The phonology of some words however points very clearly to a European Romani origin: *mar-* ‘bread’ (Romani *maro*, Domari *mana*), *šawe* ‘boy’, *čavo* and *čai* ‘girl’ (Romani *čhavo* ‘boy’, *čhave* ‘boys’, *čhaj* ‘girl’), *kam* ‘sun’ (Romani *kham*), *ker/kir* ‘house’ (Romani *kher*), *kalo* ‘black’ (Romani *kalo*), *lašo* ‘good’ (Romani *lačho*), *manuš* ‘man’ (Romani *manuš*), *rátsi* ‘night’ (Romani *rat’i*), *yag* ‘fire’ (Romani *jag*), *kagniye* ‘fowl’ (Romani *kaxni* ‘chicken’). The word *balamo/balamu* ‘Christian’ is a specifically Balkan Romani term denoting ‘Greeks’. The presence of *enna* ‘nine’ (Romani *enja*, from Greek) in Newbold’s list further confirms the Balkan Romani origin. Sampson (1928) had, on this basis, suggested that the Egyptian *Ġajar* were the descendants of Moldavian Romanies who had been taken prisoner by the Ottomans and deported to Egypt as slaves. Sampson was sceptical about some of the items on Newbold’s list, suggesting contamination with George Borrow’s lists of the Romani vocabulary of Spanish Gypsies, to which Newbold had had access. Thus, the verb *sobelar* ‘sleep’ appears in its Spanish-Gypsy (Caló) form, with a Spanish infinitive ending. However, other items on the list suggest replication of Romani inflected verb-forms, which are not present in Caló. Thus we find on Newbold’s list words spelled as *khaba* ‘eat’, *chúrabi* ‘rob’, *laba* ‘bring’, which remind us of Romani *xava* ‘I-eat’, *čorava* ‘I-rob’, *lava* ‘I-take’. The presence of inflected items is partly confirmed by Streck’s list, where we find *besheba* ‘sit

down!' (*beš-* 'to sit', possibly *bešava* 'I sit down'), and *awela* 'come here!' (*av-ela* 's/he comes'). The data thus suggest at least contacts with a community of Romani speakers, and so a Romani immigration from the Balkans to Egypt, although the circumstances of this immigration cannot be inferred from the linguistic data alone.

A further conspicuous contributor language is Domari, the full-fledged language of Gypsy groups that are scattered throughout the Middle East. Streck's (1995: 295-297) word list of the Sudanese *Bahlawān* is almost entirely derived from Domari. Many words closely resemble the citation form known to us from Jerusalem Domari (cf. Matras 1999, Macalister 1914): *tmaliyen* 'police', *kušketi* 'small', *dies* 'two', *taran* 'three', *aštar* 'four', and more. Most nouns in the list however appear in the Domari accusative form (masculine *-as*, also *-es*, feminine *-(i)a*): *qaras* 'donkey', *santas* 'dog', *kuturyes* 'European', *šunes* 'man', *šunya* 'woman', *masiya* 'meat', *jimariya* 'chicken', and more. Some inalienable possessives (body parts, kin) appear with a possessive marker. *ikyos* 'eye' (lit. 'his/her eye'). Verbs tend to appear in the Domari third person singular present form: *sutari* 'sleep', *qotari* 'steal'. This selective replication of inflected forms, both nouns and verbs, suggests that the ancestors of the present-day *Bahlawān* had access to an inflected language and so to a form of Domari that was in everyday use as a full-fledged language. It appears that the special vocabulary was retained following a shift in the community language from Domari to Arabic, a process that is well-attested in many Romani communities of Europe. Domari has also enriched the special vocabularies of other groups which were not previously Domari speakers, and we find items like *bakra* 'sheep', *gora* 'horse', *sir* 'head', *sanota* 'dog' in various special vocabularies, Arabic-based, but also in Iran and the Caucasus.

There are thus at least three pools of lexicon on which Arabic-based special vocabularies draw. The first is the indigenous, cryptolalic component, the roots of which appear to be old or even ancient, but the patterns may still be productive and allow speakers to create new lexical items. Some vocabularies draw on this source as a primary reservoir – notably the speech of the Sudanese and Egyptian *Halab* described by von Kremer (1860), Vycichl (1959) and Hanna

(1993). Others may incorporate a selection of items, apparently as a result of contact with these user groups. The second source of lexical enrichment comes from occasional contacts with other peripatetic groups and possibly also settled populations who speak another language, such as Nubians, Kurds, Romanies, or Dom, as well as with peripatetics who are users of a different special lexicon. It is through the latter type of contacts that individual items of vocabulary may be diffused far beyond the area reached by speakers of the actual contributor language. The diffusion area of some vocabulary items may therefore cover wide regions in the Middle East and Central Asia (cf. Windfuhr 2002). Finally, a third source of vocabulary items stems from the selective retention of lexicon from a former separate language spoken by earlier generations, e.g. Domari, Kurdish, or Romani. Streck (1995:302) suggested a three-way classification of special languages, based on the type and sources of vocabulary, which he labels according to the word for ‘Christian, European’, as the *xašāna*-group for the *luġa ḥalabiyya* (in which most items are internal cryptolalic formations), the *kuttur*-group for speakers of the *luġa bahlawāniyya* (containing Domari-derived items), and the *balamo*-group for the *luġa ġajariyya* (containing a significant number of words of Romani origin, and otherwise a mixed vocabulary). Although the classification is useful, the various patterns of contact among the groups and the layered vocabulary borrowings that result from them, complicate the real picture considerably.

Our final point for consideration is the Arabic influence on Domari, the archaic Indo-Aryan language spoken by populations throughout the Arabic-speaking regions and beyond, which are known by various names. Descriptions of the language which we call here Domari appear in Pott (1844-45, 1846), based on a list by Seetzen from Nablus, Palestine, in Newbold’s (1856) description of the speech of the *Kurbat* of Aleppo and Antioch and the *Duman* of Baghdad, in Groome’s (1891) list from Beirut, and in Patkanoff’s (1908) essay based on materials collected among the *Karači* of Tabriz (Iranian Azerbaijan), Maraş, and Antep (eastern Anatolia). Other Domari-speaking communities are known to exist in Lebanon, Syria, and Jordan. The most extensive documentation of the language so far is based on the Jerusalem dialect (Macalister 1914,

Matras 1999). As an archaic New Indo-Aryan language, Domari retains the Old Indo-Aryan present conjugation of the verb and passive and causative valency morphology, as well as consonantal case endings. At the same time, like Romani, it renews the past-tense conjugation in a way that is reminiscent of north-western Indian or ‘Dardic’ languages like Kashmiri, and shows, again like Romani, agglutinated case affixes. In vocabulary, Domari shows Turkish and Kurdish influences, and an immense Arabic impact, with some 50% of the Swadesh-list of ‘basic’ lexicon (in the long, 207-item version) deriving from Arabic. Arabic loans tend to keep their Arabic phonology, which means that Domari speakers, who have been bilingual in Arabic for many generations now, also have the full range of Arabic phonemes at their disposal. It is noteworthy however that some Arabic loans in Jerusalem Domari retain a pronunciation reflecting, presumably, an earlier Arabic contact variety: For example, Domari has *qahwē* ‘coffee’, cf. Jerusalem Arabic *’ahwe*. Arabic phonology partly influences the pre-Arabic or Indo-Aryan component as well. Intonation and prosody are the most strongly convergent with Arabic, and in some words pharyngealization of stops is also found in the inherited component. As in Jerusalem Arabic, the affricates *j* and *č* are undergoing a shift to *ž* and *š* respectively. Although *b* and *p* remain distinct, there is a strong tendency toward lenization of *p*. Arabic verb roots are adapted to Domari through a strategy reminiscent of most Indo-Iranian as well as Turkic languages, whereby a ‘carrier’ verb, either transitive (from *kar-* ‘to do’) or intransitive (from *hr-* ‘to become’) carries the inflection. The Arabic base that is selected is not, as in many other languages, a nominal form or *mašdar*, but a reduced form of the imperfect/imperative: *štrī-karami* ‘I buy’, *fhim-homi* ‘I understand’.

Jerusalem Domari has in effect undergone what might be referred to as structural ‘fusion’ with Arabic in the domain of clause combining. All connectors, conjunctions, interjections, and discourse markers derive from Arabic, and word order in the basic and complex clause is virtually identical to Arabic. Inflected Arabic conjunctions and particles, such as *inn-* ‘that’ or the resumptive pronoun *īyyā-*, retain their Arabic agreement inflection, leading for instance to the introduction in Domari of third-person gender distinctions which are otherwise not

present in the inherited (Indic) pronominal system. A further domain of near-complete fusion is the area of modality. Domari retains its own tense and modality inflection, but all modal and aspectual auxiliaries with the exception of *sak-* ‘to be able to’ are borrowed from Arabic, and retain, wherever relevant, Arabic person and tense inflection: *biddī laham* ‘I want to see’, *lāzim džam* ‘I must go’, *šārat rowari* ‘she began to cry’. Most sentential adverbs, as well as many temporal expressions, are Arabic, and almost the entire inventory of prepositions is borrowed from Arabic, with the exception of several person-inflected forms (such as ‘for-’, ‘about-’, and ‘with-me, you, etc.’): *ma* ‘with’, *la* ‘to’, *fī* ‘in’, *bēn* ‘between’, *min* ‘from’, *ind* ‘at’, etc. While Macalister’s (1914) description of Jerusalem Domari still shows the full Indic series of numerals, present-day Domari (Matras 1999) has retained only the Indic numerals for 1-5, 10, and 100, replacing the others by Arabic numerals. A lexical-typological oddity is the wholesale borrowing of the comparative-superlative form and with it the Arabic lexical form of the adjective, so that all Domari adjectives, even the basic adjectives that are Indic, have suppletive, Arabic-derived non-positive forms: *tilla* ‘big’ – *’akbar* ‘bigger’, *kištota* ‘small’, *’azyar* ‘smaller’.

Finally, Domari is undergoing further convergence with Arabic by generalizing those structures that are closest to the counterpart Arabic configuration. In the domain of adjective attribution, the inherited Indic word order is Adjective-Noun: *tilla zara* ‘the big boy’, *tillī šōnī* ‘the big girl’. However, there is a clear preference toward the use of predicative adjectival constructions in place of the normal attributive construction, for the former agree in their word order with the Arabic pattern. Thus, we normally find *zarēk tillēk* ‘the big boy (=the boy, being big)’, *šōnik tillik* ‘the big girl (=the girl, being big)’. In the possessive attributive construction, the formation Possessor-Head (*bāyim kuryos* ‘my father’s house’, *barim kuryos* ‘my brother’s house’) is being replaced by the construction Head-Possessor: *kury-os bāyim-ki* ‘my father’s house’, literally ‘his-house of-my-father’, cf. Arabic *bēt-o la-’abūy*. Although word order in the verb phrase is generally identical to Arabic, Domari has retained just one trace of the Indic verb-final order, namely the present-tense enclitic copula, which in Arabic is matched by the nominal clause: *ama mišta-hromi* ‘I am ill’, *pandži mišt-ēk* ‘he

is ill', cf. Arabic *'ana marīd-Ø, huwwe marīd-Ø*. In the other tenses, Arabic auxiliaries are employed to maintain the similarity to the Arabic construction: *ama kunt mišta-hroma* 'I was ill', *pandži kān mišt-ēk* 'he was ill', cf. Arabic *'ana kunt marīd, huwwe kān marīd*. It is perhaps useful to enumerate those areas of structure that are not prone to convergence with, or substitution through, Arabic: They include a selection of basic vocabulary, perhaps as few as 500-600 lexical roots, among them most body parts and verbs of movement and physical activity, some but not all kinship terms, and the numerals under 5; gender, number and person agreement rules, and corresponding nominal and verbal (and to a lesser extent adjectival) inflectional morphology, including subject and object concord and possessive inflection; synthetic valency-derivation, aspect, tense, and modality formation in the verb; pronouns and demonstratives, as well as place deixis and some time deixis expressions; some basic expressions for local and spatial relations, including some inflected prepositions that derive from them. Arguably, this is an extremely limited set of structural features, and a limited vocabulary range. The primary function of such a system appears to be to maintain and flag group separateness, which makes it functionally related, albeit only partially, to the special vocabularies discussed above.

School of Languages, Linguistics, and Cultures
University of Manchester, Manchester M13 9PL, UK
 yaron.matras@manchester.ac.uk

References

- Burton, Richard F. 1898. *The Jew, the Gypsy, and El Islam*. London: Hutchinson & Co.
- Canova, Giovanni. 1981. "Notte sulle tradizioni zingare in Egitto attraverso la testimonianza di un capo Nawar". *Lacio Drom* 17/6.4-25.
- Finck, Franz N. 1907. "Die Grundzüge des armenisch-zigeunerischen Sprachbaus". *Journal of the Gypsy Lore Society*, New Series, 1.34-60.

- Galtier, Émile. 1912. Mémoires et fragments inédits. *Mémoires publiés par les membres de l'Institut Français d'Archéologie Orientale de Caire* 27.1-9.
- Hanna, Nabil Sobhi. 1993. *Die Ghajar*. Munich: Trickster.
- Kruse, Fr. 1854. *Ulrich Jasper Seetzen's Reisen durch Syrien, Palästina, Phönicien* [...]. II. Berlin: Reimer.
- Littmann, Enno. 1920. *Zigeuner-Arabisch. Wortschatz und Grammatik der arabischen Bestandteile in den morgenländischen Zigeunersprachen*. Bonn & Leipzig: Kurt Schroeder Verlag.
- Lorimer, D. L. R. 1939. *The Dumāki Language*. Nijmegen: Dekker & van de Vegt.
- Macalister, R. A. S. 1914. *The language of the Nawar or Zutt, the nomad smiths of Palestine*. (Gypsy Lore Society Monographs 3). Edinburgh: Edinburgh University Press.
- Matras, Yaron. 1999. "The state of present-day Domari in Jerusalem". *Mediterranean Language Review* 11.1-58.
- Matras, Yaron. 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge 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.
- Patkanoff, K.P. 1908. "Some words on the dialects of the Transcaucasian Gypsies – Boša and Karaci". *Journal of the Gypsy Lore Society*, New Series, 1.229-257; 2.246-266, 2.325-334.
- Pott, August F. 1844-1845. *Die Zigeuner in Europa und Asien*. Halle: Heynemann.
- Pott, August F. 1846. "Über die Sprache der Zigeuner in Syrien". *Zeitschrift für die Wissenschaft der Sprache* 1.175-186.
- Rao, Aparna. 1995. "Marginality and language use: The example of peripatetics in Afghanistan". *Journal of the Gypsy Lore Society*, Fifth Series, 5.69-95.
- Sampson, John. 1928. "The Ghagar of Egypt: A chapter in the history of Gypsy migration". *Journal of the Gypsy Lore Society* Third Series 7.78-90.
- Streck, Bernhard. 1996. *Die Ḥalab*. Munich: Trickster.

- Vychichl, Werner. 1959. "The slang of the Ḥalab is-Sūdān". *Journal of the Sudan Antiquities Service* 7.223-228.
- Windfuhr, Gernot. 2002. "Gypsy dialects". *Encyclopædia Iranica* ed. by Ehsan Yarshater, Volume 11, 415b-421b. London & New York: Routledge & Kegan Paul.
- Winkler, Hans A. 1936. *Ägyptische Volkskunde*. Stuttgart: Kohlhammer.