

The State of Present-day Domari in Jerusalem

Author(s): Yaron Matras

Source: *Mediterranean Language Review*, Vol. 11 (1999), pp. 1-58

Published by: Harrassowitz Verlag

Stable URL: <https://www.jstor.org/stable/10.13173/medilangrevi.11.1999.0001>

## REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/10.13173/medilangrevi.11.1999.0001?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/10.13173/medilangrevi.11.1999.0001?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



Harrassowitz Verlag is collaborating with JSTOR to digitize, preserve and extend access to *Mediterranean Language Review*

JSTOR

## THE STATE OF PRESENT-DAY DOMARI IN JERUSALEM

### 1. INTRODUCTION

Domari is an endangered Indic language spoken in a socially isolated and marginalised community of formerly itinerant metalworkers and entertainers in the Old City of Jerusalem. Speakers refer to themselves in their language as *dōm* (singular & collective) or *dōme* (plural). The term *nāwar* ('Gypsies') is adopted quite freely in Arabic conversation, although it is disliked due to its derogatory connotations. Domari is part of the phenomenon of Indic diaspora languages spoken by what appear to be descendants of itinerant castes of artisans and entertainers who are spread throughout Central Asia, the Near East, and Europe. They include rather loosely related languages such as Dumaki (Hunza valley in northern Pakistan; Lorimer 1939), Parya (Tajikistan; Oranskij 1977, Payne 1997), Lomavren (Indic vocabulary in an Armenian grammatical framework; Finck 1907, Patkanoff 1907/1908), Inku (Afghanistan; Rao 1995), and Romani (primarily Europe and Asia Minor), the latter being by far the most widely documented and the most thoroughly described. Apart from Jerusalem and the West Bank, Domari-speaking communities are known to exist today in Gaza, Jordan, Lebanon, and Syria. Here too, the language is reported to be in rapid decline and is apparently in active use only among the older generation. Remnants of closely related idioms survive in secret lexicons employed by itinerant communities elsewhere in the Near East; such contemporary usage of Domari-based lexicon has been documented for the Kurdish-speaking Mıtrıp or Karaçi commercial nomads of eastern Anatolia (Benninghaus 1991) as well as for the Luri speech of the Luti people of Luristan (Amanolahi & Norbeck 1975). The latter, along with other samples of lexically related varieties such as those published by Gobineau (1857), Patkanoff (1907/1908), and others, provide evidence that Domari dialects were once spread between western Iran, eastern Turkey, and southern Palestine, spoken by groups known mainly by the names Dom, Kurbati, and Karaçi.

The question of the affiliation of Domari has occupied Romani linguistics ever since the first appearance of Domari wordlists, and later of fragmented grammatical sketches, in the 19th century. Seetzen's list of some

300 lexical items collected among nomadic Dom in the vicinity of Nablus was made available to scholars before its actual publication in a travel diary edited by Kruse (1854) and was incorporated into Pott's (1844–1845) monumental comparative overview of Gypsy dialects. Pott (1846) and Paspati (1870) both published further wordlists and notes on grammatical structures based on sources from Lebanon and eastern Anatolia respectively. Newbold's (1856) wordlists documented the use of Domari among the Kurbáti of northern Syria and the Dumán of Baghdad, Groome (1891) published samples from Beirut and Damascus, and Patkanoff (1907/1908) presented material that appears to have originated from Azerbaijan. It is on the basis of the material from these sources that Sampson (1923) postulated a single origin for Romani and Domari (as well as Lomavren) and a common migration from India, suggesting that a split took place in Persian territory. Sampson referred to parallels in the noun and verb inflection to justify his theory. Drawing on separate phonological developments – notably reflexes of aspirated voiced stops – he coined the terms *phen-Gypsy* (Romani and Lomavren) and *ben-Gypsy* (Domari) for the two branches (< Old Indo-Aryan *bhagin-* 'sister'). It has since been widely accepted (cf. Hancock 1988) that all three ethnonyms – Dom, Rom, Lom – are derived from the Indic *dom*, a caste name, although their origin in a low-caste of marginalised and stigmatised service-providers of various kinds has more recently been contested (Hancock 1998). While the exact historical connection between Domari and Romani remains unclear, modern studies in Romani linguistics acknowledge at least the possibility of a close link (see e.g. Hancock 1988, Bubeník 1997, Boretzky 1995; but see Hancock 1995 for a critical view).

Jerusalem Domari has been made known to the academic world by Macalister (1914), whose monograph, a reprint of a series of articles published in the *Journal of the Gypsy Lore Society* between 1909–1913, constitutes the only attempt to date to provide a full description of any Domari-type variety. Macalister worked with only one informant, but was able to compile a glossary comprising over 1000 entries, as well as a collection of over a hundred short texts (many of them however translations of material provided by Macalister himself and not authentic stories). Macalister's data have since served as the source for most if not indeed all linguistic discussions of Domari, including specialised investigations devoted to its Arabic and Iranian loan components by Littmann (1920) and by Barr (1943) respectively.

First-hand linguistic descriptions of Domari proper (i.e. excluding secret lexical insertions) following Macalister (1914) are virtually non-existent: Yaniv (1980) provides an introductory ethnographic outline of the Dom community in Jerusalem and the West Bank, in which he includes several proverbs in Domari, and Nicholson, in an unpublished and undated manuscript, glosses and interprets a six-line transcript of a message in Domari recorded in Syria.<sup>1</sup> All other work on Domari (e.g. Kenrick 1975–1979) is based on replications and discussions of earlier sources. The present contribution thus enters a void that has been left since the publication of Macalister's work on Domari nearly a century ago. It is based on elicited questionnaire material, narratives, and conversations recorded between 1996–1999 in the Old City of Jerusalem, in the very community whose language was the subject of Macalister's study.<sup>2</sup>

Due to the limited space available in the present format, it is not my intention to provide a detailed and comprehensive description that would overlap with, or succeed the one published by Macalister. Rather, my aim is to provide a concise overview of the grammatical structures of Domari highlighting especially those aspects that were not noted or discussed by my predecessor Domarist. Following a brief description of the community I discuss features of the Domari sound system and morphosyntax. I then go on to an assessment of the Arabic component, followed by observations on linguistic stratification and language decline. I conclude with a note on the linguistic-typological affinity between Domari and Romani. With the exception of citations from Macalister that are inserted for the sake of comparison, and which are always marked as such, all data presented below derive from my own corpus of recordings.

- 1 Sample sentences of Syrian Domari were also collected, transcribed, and privately circulated by Marielle Danbakli, 1998.
- 2 A total of around two dozen speakers and semi-speakers were interviewed. I wish to express my gratitude to the residents of Burġ el-Laqlaq, and in particular to the Sleem families, for their cooperation. I also wish to thank the following people for their support and encouragement, and for stimulating discussions about issues dealt with in this contribution: Peter Bakker, Donald Kenrick, Viktor Elšík, Yigal Tamir, Tom Gross, Gilad Margalit, Victor Friedman, Jonathan Freud, Amoun Sleem, Miron Benvenisti. I am alone responsible for the views expressed here.

## 2. ETHNOGRAPHIC BACKGROUND

The bulk of the Dom population of Jerusalem lives in rented accommodation in the northeastern corner of the Muslim quarter of the Old City, just north of Lions Gate and the northern entrance to the Ḥáram, the compound which includes the Dome of the Rock and Al-'Aqṣa Mosques. Some families have left the Old City in recent years to settle in the neighbourhoods, villages, and suburbs in and around East Jerusalem. A sizeable community of expatriates lives in Amman, Jordan, having fled Jerusalem and the West Bank during the Six-Day War in June 1967. Many Dom maintain family ties with the Amman community, travel to Jordan regularly, and host visiting relatives in Jerusalem. There are no reliable figures about the size of the Jerusalem Dom community. Members of the community claim a total population of up to one thousand, a figure which is accepted by some observers. A survey carried out in the mid-1970s by an Israeli anthropologist put the entire Dom population of the Old City at the time at between 200-300,<sup>3</sup> which coincides with the figure of 300 given by Yaniv (1980). This might suggest a number of only around 600-700 today.

The Dom are Sunni Muslims and live among Palestinian Arabs, with whom they share cultural traditions, infrastructure such as accommodation, education, and services, and a variety of everyday concerns. Arabic has now become the principal language of the Dom community. I estimate that only around 20% of adult Dom use Domari as the language of daily interaction in their homes; the great majority of them are over fifty years of age. Among the remainder there exists a sizeable group with passive knowledge of the language, to varying degrees, while the younger generation under twenty years of age is familiar at best with a small inventory of words. Only a few elderly speakers could be encountered who used or were at all familiar with the designation *dōmari*, cited by Macalister as the name for the language and replicated since in linguistic literature as such. Most members of the community, including active and fluent speakers, used the term *dōm*, and occasionally an Arabicised version *dōmī*, to refer both to the people and the language. The younger generation, whose knowledge of the language is fragmented, claimed never to have even heard the term *dōmari*.

The language itself maintains internal designations for a number of ethnic groups. This is rather typical of cryptic word-formations in in-group

3 Yigal Tamir, personal communication, 1998.

idioms of isolated, non-territorial or peripatetic communities, which tend to construct ethnonyms using native materials. In Domari they include *ṭāt* '(settled) Arab',<sup>4</sup> *qrara* 'Beduin, Jordanian', *džudži* 'Egyptian', *ktīr* 'Christian', *portkīla* 'Jew', *nohra* (= 'red') 'British'. Some of those may be recycled names for peoples who were contiguous with the Dom before their immigration into Palestine, others are perhaps borrowings from other languages. The term *nohra* is clearly a purposeful creation inspired either by the red caps worn by British forces, or else by the red shade of pale skin exposed to the Mideastern sun. As for *portkīla*, a connection has been suggested to the citrus-growing enterprises of Jewish settlers in the coastal plain area, based on Arabic *burṭūqāl* 'orange' (< *Portugal*);<sup>5</sup> an Arabic-based etymology however would not explain the reconstruction of initial *p* in Domari, but, since the term is already attested by Macalister, Turkish *portakal* is a possibility.

The origin of the Jerusalem Dom is in a group of commercial nomads. The immediate ancestors of today's community were tent-dwelling smiths and tanners who settled within the boundaries of the Old City walls in several waves beginning in the 1940s, and lasting until after the Israeli occupation in 1967, when the last Dom abandoned their tents and moved into permanent dwellings. Begging was apparently practised by the women in the community until shortly after the Israeli occupation. It is still practised in the city by expatriate Jerusalemites based in Amman, who return to Jerusalem during the Muslim holiday seasons and can be met begging around the entrances to the main Mosque compound. Young Dom women, accompanied by small children, can also be seen begging at various times of the year in other parts of the Old City, mainly around Jaffa Gate, which is the principal access route used by tourists, and the Muristan market which is adjoined to the Church of the Holy Sepulchre, a major tourist attraction. Although some of these women are residents of the West Bank, most are Egyptian citizens from Al-'Arish in the northern Sinai who enter Israel on a tourist visa. Those interviewed by me were semi-speakers of Domari, with a retrievable active knowledge of only some words or phrases.

4 I am grateful to Marcel Erdal for pointing out that this word was used in Mamluk Turkic in the meaning 'Arab, fallah' from the 14th century on.

5 Yigal Tamir, personal communication, 1998.

The Jerusalem Dom typically distance themselves from these visitors and emphasise that the local Dom community has no part whatsoever in begging activities. There appear to be two factors underlying this attitude. The first is to be seen in a deeply-rooted socio-cultural gap between the Dom of Gaza and Al-'Arish (formerly also of Jaffa), whose traditional occupations included musicians, dancers, and other forms of entertainment such as bear- and monkey-leaders, and the artisans and craftsmen who settled in Jerusalem. Although the Jerusalem Dom have also produced their share of musicians, who are the community's pride, occupations such as table-dancing, though still common in the 1940s, are now strictly forbidden by the community.

The second explanation for the distance between the two groups is the relative economic stability and standard of living and education which the Jerusalem Dom enjoy, compared to the other communities in Gaza, Egypt, or Jordan. The great majority of the economically active community members have been in paid employment for several decades now. Already during the 1940s, under British rule, Dom began working for the Jerusalem municipality, often in trash disposal and as caretakers in public lavatories. Under Jordanian rule, it is claimed, the municipality's environmental health department depended largely on labour from among the Dom community. This state of affairs was then inherited by the Israeli authorities following the occupation in 1967. This first generation of manual labourers paved the community's way out of traditional crafts and peripatetic services and into paid employment and rented accommodation. The younger generation now seeks paid employment in a much broader range of sectors. Since the municipal annexation of East Jerusalem to Israel in 1967, the Dom have access to the Israeli system of welfare and state benefits, including pension, child benefits, public health care, and free primary education. Today's young generation Dom, and in particular the women in the community, are the first to have enjoyed unrestricted access to elementary education. Moreover, they are the first to continue into specialised vocational training, especially nursing. The changing socio-economic profile of the community has deepened the gaps between the Jerusalem Dom and those in other parts of the country. It is also fracturing the traditional overlap between ethnic identity and socio-economic identity which had existed previously among the Dom, giving rise to a feeling of an ethnicity vacuum.

Despite their immersion into Muslim Arab society, the Dom nevertheless maintain their awareness as a separate ethnic entity, partly by tradition,

and partly as a result of continuing marginalisation and isolation. All members of the community, especially the younger generations, report incidents of racist abuse and discrimination by mainstream Palestinian society. As a result, many highlight their indifference to Palestinian political aspirations. On the other hand they are regarded by Israeli authorities as an integral part of the Arab population of East Jerusalem and the West Bank, and suffer the same occasional if not continuous abuse from them. Politically this positions the Dom in actual fact much closer to Palestinian than to Israeli society. The Jerusalem Dom have also had occasional contact with European Roma who have visited their community, usually as missionaries for various religious movements. This exposure has prompted a sense of curiosity toward the Romani community in Europe, and a number of young Dom have in recent years been exploring the prospects of developing community institutions modelled in some way or another on the experience of Romani cultural and political associations in Europe and the United States.

A traditional link appears to exist between the Dom community and another group of itinerant metalworkers of the West Bank, known as 'Kurds' (Arabic *krād*, Domari *krāde*) or, reportedly, in their own speech as *rōm* or *rōmāt*. The two groups intermarry, and to some extent are familiar with each other's languages. The speech of the Kurds however appears to be a secret lexicon consisting of items from Kurdish as well as Domari and probably of other sources as well. Their indigenous name suggests a connection to the Rom of Europe, and such a connection cannot be ruled out given the presence of Romani items in the secret vocabularies of other itinerant groups in the Near East, such as the Ghagar of Egypt (Newbold 1856) or the Poşa of eastern Anatolia (Benninghaus 1991).

Apart from their language, traditional occupations, dress, and songs, all of which are disappearing or have already disappeared from Dom community life, today's Dom are able to point to few traditions that distinguish them from mainstream Muslim Palestinian-Arab society. The most important feature cited in connection with customs is the pilgrimage to Nabi Mūsa, according to Muslim tradition the burial place of Moses, in the Judean desert. Although gatherings at this site are common to all Muslims of the region, the Dom have a specific day, in April, on which they gather there. One may wish to draw parallels to the importance of pilgrimage in European Romani culture. An additional cultural feature which is reminiscent of Romani traditions is the existence of tales explaining the group's



destiny of wandering as punishment for an ancient sin. Romani tradition has the story of the nails stolen from the cross, in numerous variants (see e.g. Pickett & Agogino 1960). Among the Dom, two distinct tales can be heard, which on occasion are combined into one.<sup>6</sup> The first portrays the Dom as descendants of Džassās, the leader of the tribe of Banū-Marra. In a conflict with another tribe, Džassās killed the rival leader, Klēb. He was then killed in revenge by Klēb's younger brother, Sālem ez-Zīr. Ez-Zīr continued to persecute Banū-Marra and drove them into exile, ruling that they should no longer be allowed to ride horses but only donkeys, that they should remain outdoors, and that they should not be allowed to stay in one place for longer than three nights. In slight variation, the story of ez-Zīr is documented for other itinerant cultures in the Near East, notably the Egyptian Ghagar and Nawar, by Newbold (1856: 291) and Canova (1981). Like the Ghagar of Egypt, as reported by Newbold, the Jerusalem Dom too assert the existence of a written document in which the story of their origin is described.

A second tale depicts the Dom as descendants of a nomadic tribe of entertainers who were settled in Iran, were given farmland and animals and expected to become farmers. The king, who had invited them to settle, later discovered that they had neglected their lands and instead spent most of their time singing and dancing. He then banished them from his kingdom and they became nomads again. This story is obviously related to the story of Bahram Gur as told in Firdusi's *Shāh Nāme*, and indeed the Mukhtar of the Jerusalem Dom community cites Bahram Gur as the name of the king in his story. Apparently, as pointed out by Yaniv (1980), the legend is not an authentic Dom story but was adopted more recently from foreign sources, through indirect exposure to the literature on Gypsies. It is noteworthy however that Amanolahi & Norbeck (1975: 3) report a similar story of descent which is told about the Luti of Luristan.

From this it seems that the Dom community has long been able to sustain a distinct identity through a delicate balance of language maintenance, integration into a regional context of peripatetic cultures, and macro-level cultural integration into sedentary Arab society. The latter has become by far the dominant feature of Dom culture in recent decades. It

6 Both stories already appear in Yaniv (1980). During fieldwork in Jerusalem I was also able to record them in several variations and from more than one speaker.

coincides with the gradual decline of nomadism and traditional service occupations as well as with the decline of the Domari language.

### 3. THE DOMARI SOUND SYSTEM

For practical purposes the present discussion will employ a simplified, rather than a narrow transliteration of sounds in the sections dealing with morphosyntactic structures. This notation does not pretend to convey a system of phonemic oppositions in the strict sense, i.e. to depict only potentially meaningful sound distinctions. Rather, it is a compromise system, which takes for granted the variability of sounds. The present section outlines patterns of variation and takes an inventory of individual sounds encountered in the corpus.

#### 3.1 Vowels

A simplified rendering of the Domari vowel system might convey a rather symmetrical picture of five cardinal vowel phonemes /a e i o u/ and a straightforward opposition of +/- length. Macalister however, while accepting a five-term system for long vowels – described using English and French examples roughly as [æ:, e:, i:, o:, u:] – suggests a more complex inventory of short vowels, with [æ, a, ε, i, ɔ, Δ, u]. This impression of a stronger diversity among short vowels can be supported here, though more articulations are recognised than in Macalister's description (see Figure 1). Within what can be considered the *a* group of open vowels, three positions are identified: Middle [a] as in [ma'nʊs] 'person' or [ʔa'ha] 'this (m.sg)' occurs consistently only in a comparatively small number of recorded items, and appears elsewhere to be in free variation with semi-fronted [æ], as in [næn'dom] 'I brought', [læ'ʃi:] 'girl'. When both variants co-occur in a word, [a] is typically in stressed position, [æ] is unstressed: [bæ'nam] 'I shut (subjunctive)'. Alternation is also sensitive to open and closed syllable position: [mam] 'uncle', [mæ'mi:] 'aunt'. Back [ɑ] tends to occur around pharyngealised consonants, trills, and semi-vowels: [dɑnd] 'tooth', [prɑ'na] 'white', [dɑ'wɪʃ] 'dance'. [Δ] is rare, and appears most consistently in ['pΔndʒɪs] 'five', otherwise in paradigmatic relation to, and in variation with *a*-vowels: [gɑr'dɪk, gar'dɪk] '(she is) well', [laka'dom, laka'dom, lake'dom] 'I saw'. Among the *e*-vowels, closed [e] as in [ʔe'me] 'we' appears alongside open [ε], most salient in the plural ending – [mɑ'ʔε] 'people'. Among the *i*-vowels, closed [i] is the only sound that appears in final position – [go'di] 'wit', semi-open [ɪ] is confined to pre-consonantal

positions – [tɪˈlla] ‘big’, as is the case with centralised [ɨ] – [dʒɨˈwɪr] ‘woman’, [gɨf] ‘all’. This distribution is paralleled by the *u*-vowels, with [u] in [ɑˈtu] ‘you’, [ʃukˈna] ‘oil’, alongside semi-open and centralised variants in pre-consonantal position as in [pʊˈtər] ‘son’. Short [o] often alternates with its long version; consistent short pronunciation is found in [ʔaˈdʒoti] ‘today’. [ɔ] is rare, appearing in [dɔˈwami] ‘I wash’ (cf. [daˈwami] ‘I dance’).

Figure 1: Inventory of vowels

i i:	ɨ	ʊ	u u:
ɪ			ʊ
e e:			o o:
ɛ		ʌ	ɔ
æ æ:			
a a:		ɑ ɑ:	

As mentioned above and shown in Figure 1, short vowels are more diverse in quality than long vowels. Interchangeability is common among adjacent articulatory positions, the most common interchangeable pairs being [a-æ], [a-ɑ], [a-ʌ], [ʊ-ɨ], [ʊ-ɪ], [o-ʊ], [e-ɛ], [ɛ-æ], [e:-i:], [o:-u:]. Such variation is often the product of regressive assimilation triggered by distinct grammatical endings: [wʊˈda] ‘old man’, [wɨˈdi:] ‘old woman’. Variation among adjacent vowel positions, partial centralisation of high vowels and the fronting of raised back vowels [u > ʊ > ɨ; u > ʊ > ɪ] are processes that are shared with Palestinian as well as with northern Levantine Arabic, and with Kurdish – all significant contiguous languages in the recent history of Domari. From among the overall inventory of vowel sounds, only [ɔ] and [ʌ], both rather infrequent, are not shared with local Palestinian Arabic. Prothetic and epenthetic vocalisation around consonant clusters may also be regarded as a regional phenomenon.

Length is characterised by a tonal lengthening of the vowel, best recognisable in first syllable position in bi- and multi-syllabic words: [doːˈmɛ] ‘Doms’, [ʔuːˈjar] ‘town’. A rare minimal pair is [ʔɔʔ] ‘sun, heat’, and [ʔɑʔ] ‘sedentary Arab, Fallah’, confirming nonetheless the distinctive function of length opposition. Length is often neutralised, however, usually in final position, as well as in pre-final positions in grammatical endings, showing

[e:>e] and [i:>i]: [a'h're:ni:] 'we are', alongside [a'h'reni]. Among the long vowels, only the *ā*-vowels show variation in quality: [ba<sup>ad</sup>] 'grandfather', [ta<sup>a</sup>t] 'sedentary Arab, Fallah', [wæ<sup>æ</sup>j] 'wind'.

### 3.2 Consonants

Here too there is a tendency towards a merger with Arabic, evident both in the incorporation of Arabic lexical loans without any obligatory phonological adaptation, and so in the wholesale accommodation of Arabic phonemes, as well as in the infiltration of Arabic sounds into the inherited (non-Arabic) component (see Figure 2).

Perhaps the most conspicuous feature is the pharyngalisation of dentals, which is distinctive within the Arabic component, but to a large extent variable within the Indic or pre-Arabic component. We thus have the alternations [do:m, ɖo:m] 'Dom', [tat, tət] 'heat', [mu'tur, mu't̪ur] 'urine'. Conventionalisation of pharyngealisation in non-Arabic items can be found in the tendency towards progressive assimilation, where a Domari ending follows an Arabic stem, as in [ta:w'le:ta] 'on the table', Arabic *tāwle* and Domari dative ending *-ta*. There are in addition quite a few non-Arabic lexical items which seem to have adopted pharyngealisation and which display it consistently; examples are [ɖaɳɖ] 'tooth', [maɳ] 'person', [waɳ] 'stone'.

Figure 2: Inventory of consonants

p	t	t̪	(tʃ)	k	q		
b	d	ɖ	dʒ	g			ʔ
m	n			ŋ			
	l			ɭ			
	r						
f	s	ʂ	ʃ	x	(χ)	ħ	h
(v)	z	ʐ	ʒ	ɣ		ʕ	
w			j				

Gemination is a further phenomenon that is typical of the Arabic component – [hɪ'bbo:mi] 'I like' – though stem gemination also occurs sporadically in inherited lexical items: [tɪ'lla] 'big', [ka'ʒʒa] '(non-Dom) man'. More widespread distinctive gemination can also be found within the inherited component where it is the result of consonant assimilation at the

attachment point of grammatical affixes: [xɪznawɪ'de:ssan] < *xiznawidēs-san* 'you (pl.) made them laugh', but [xɪznawɪ'de:sa] < *xiznawidē-san* 'they made them laugh'; [la'harrɪ] < *lahar-r-i* 'he sees you', but [la'hari] < *lahar-i* 'he sees', [kur'jamma] < *kuriya(n)-ma* 'in the houses', but [kur'jama] < *kuriya-ma* 'in the house'.

The pharyngals [ħ] and [ʕ] appear to be restricted to the Arabic component. There are other consonants that may be assigned predominantly but not exclusively to Arabic loan material. Thus [ɣ] appears occasionally in pre-Arabic items, as in [je'ɣer] 'horse', [biɣ] 'moustache',<sup>7</sup> as does [q] – [qɑ'jɪf] 'food', [qo:l'dom] 'I opened', alternating frequently with [k]: [ka'pi, qa'pi] 'door' (< Turkish *kapı*), [kaʃto'ta, qaʃto'ta] 'small'. [q] is further subject to variation with [χ], as in [qo:l'dom, χo:l'dom] 'I opened', [qal, χal] 'said' (discourse particle introducing quotations in narratives, from Arabic *qāl* 'he said'). The realisation in Domari of [q] in Arabic-derived words such as ['qahwa] 'coffee' points to an early adoption of this component and to its current perception as an integral part of the Domari system. When conversing in Arabic, Doms will consistently adopt the Jerusalemite pronunciation ['ʔahwɛ]. The etymological Arabic consonants [θ] and [ð] however do not appear in the material, and their contemporary Palestinian Arabic reflexes [t] and [d, z] are found instead. A further consonant that is typical of the Arabic lexical component is [ʔ], though it also functions regularly within the pre-Arabic component indicating verb negation in final position: [bi:'reʔ] 's/he does not fear'.

Incongruent with the contiguous Arabic system are the sounds [p], [g] (found in Egyptian, but not in Palestinian Arabic), [ŋ], as well as [tʃ] (found in rural dialects of Arabic in the regions surrounding Jerusalem to the west and northwest, an outcome of palatalisation of underlying [k]: *čalb* < *kalb* 'dog'). All four are restricted to the pre-Arabic component: [pi'rɪn] 'nose', [gur'gi:] 'throat', [maŋ'ami] 'I like', [tʃan'tʃɪmma] 'next to me' (*čanč-i-m-ma* 'in my vicinity'). Although the [p-b] contrast remains on the whole distinctive – cf. [pa'jo:m] 'my husband', [ba'jo:m] 'my wife' – there are signs of its partial retreat. In initial position, [p] often undergoes lenition: [pʊ'tər] 'son', [pʌndʒi] 's/he'. In medial position, fricativisation can be observed: [kafi'ja] 'door (acc.)'. Also contrasting with Arabic we find, though marginally in the corpus, a voiced labio-dental fricative [v], in variation with [w]: [rov'rom] alongside [row'rom] 'I wept', occasionally

7 Here the sound [ɣ] is borrowed along with the Turkic word.

replacing original Arabic [w] as in [ʁv'ldrom] 'I was born'. A velar /-sound also appears rather marginally in [ʃaʔ] 'well, waterhole', [saʔ] 'rice'; it is shared with southern Kurdish, and with Arabic in *'allah, yallah*, and in the environment of pharyngals, as in *xallaʃ*.

A case of sound convergence with Arabic is the status of the alveopalatal affricates [dʒ] and [tʃ]. The first exists in principle in the inherited inventory of Palestinian Arabic, but is undergoing reduction to a simple fricative [ʒ]. This process is reflected in Domari as well; a general retreat of affricates becomes apparent when one compares our material with that discussed by Macalister. Some words tend to maintain the underlying affricate rather consistently: [la'dʒi] 'shame' [dʒu'dʒi] 'Egyptian'. Affricates are also generally retained following dentals: [pandʒi] 's/he', [man'dʒa] 'inside'. Elsewhere, there is variation, and in pre-consonantal position, general reduction: [dʒa, ʒa] 'go', [xu'dʒoti, xu'ʒoti] 'yesterday, [dʒib, ʒib] 'tongue', but [ʒbo:m] 'my tongue'. The voiceless counterpart, which lacks an Arabic match, undergoes a similar though more radical change. In the speech of most speakers interviewed, the affricate is, except among the oldest speakers, almost entirely lost in initial position – [ʃir'dom] 'I spoke' < *čirdom*, [ʃo:'ni:] 'girl' < *čōnī* – and subject to variation in medial position: [la'tʃi:, la'ʃi:] < *lāčī* 'girl'.

### 3.3 Stress

Domari has word-level stress, contrasting with the Arabic phoneme-level stress (with accented long vowels). Stress falls on the last syllable of lexical items (*ūyár* 'town'), as well as on the grammatical markers for gender/number (*šōn-ál/šōn-é* 'boy/boys'), Layer I case inflection (see below; *dōm/dōm-ás* 'Dom.nom/acc.'), possessive personal markers on the noun (*bāy-óm* 'my father'), person inflection in prepositions (*atnī-r* 'about you'), subject concord markers on the verb (*lahed-óm* 'I saw'), and the postposed synthetic negation marker on the verb (*inmangam-é* 'I don't like'). Unstressed grammatical markers are Layer II case markers (*ūyár-ma* 'in the town'), tense markers (see below; *lahedóm-a* 'I had seen'), and enclitic object markers (*lahedóm-ir* 'I saw you'). In these accentuation patterns Domari, disregarding its particular forms of enclitic object and possessive personal markers and of synthetic negation, matches exactly the features of accentuation encountered in conservative dialects of European Romani. The most noticeable difference between the two languages is the treatment of recent loan nouns. In Romani, European-origin nouns usually have non-ultimate stress in nominative forms. In Domari, Arabic nouns are

adapted to ultimate accentuation patterns: *baladiyyá* ‘municipality’ < Arabic *baladíyya*. Exceptions are proper nouns, which retain their original stress in the nominative form – *áḥmad* – but adapt in inflected forms – *aḥmadás* (acc.).

#### 4. NOUN PHRASE MORPHOLOGY

##### 4.1 Definiteness and indefiniteness

Domari has no definite article; the construction referred to by Macalister (p. 8) as the ‘superdefinite article’, consisting of a structurally reduced demonstrative that is attached to the noun and lacks contrastive deictic function, could so far not be found in the corpus. The only feature that might resemble the structure described by Macalister is the use of a full adnominal demonstrative with a reduced contrastive deictic function; consider the following excerpt from a story, and the literal translation that follows:

- (1) a. *qrara aha šōna, mangida ihī dōmiyē*  
 Beduin this boy asked.3sg this Gypsy.girl.acc  
*min bāyiskī.*  
 from father.poss.abl
- b. *rfaḍkeda aha bāyos aha čōnaskī.*  
 refused.3sg this father.poss this boy.abl
- c. *tānī dīs aha šōna qrara, putros šēxaskī, gara ta*  
 second day this boy Beduin son.poss sheikh.abl went.3sg comp  
*lakar dōman mā lakedosan, rawirdedi min*  
 see.3sg.subj Gypsies.acc not saw.3sg.3.pl departed.3pl.perf from  
*hundar min uhu dēsoskī.*  
 there from that town.poss.abl
- d. *ehe dōme rawarde gare krēn? šīrāqata*  
 these Gypsies left.3pl went.3pl where Iraq.dat
- e. *aha šōna šār rōwari atnīsan.*  
 this boy began.3sg cry.3sg about.3pl
- a. The Beduin, this boy, asked his father for this Dom girl.  
 b. This father of this boy refused.  
 c. The next day this boy the Beduin, the Sheikh’s son, went to find the Doms but he didn’t find them, they have departed from there, from that town.

- d. Where did these Doms go? To Iraq.
- e. This boy began to cry for them.

Quite clearly Domari has rather lenient rules for the insertion of adnominal demonstratives, as far as deictic focus is concerned: They do not necessarily convey a shift in focus, nor are they confined to new topical entities or to the retrieval of topics from earlier, rather than immediately adjoined portions of the discourse. It seems more appropriate nevertheless to speak of a reduced contrastive-deictic function, rather than of its complete loss in this connection. Thus it seems clear that even the functional criteria for the grammaticalisation of demonstratives as definite articles (cf. Diessel 1999: 19–20) are not entirely met, while structural criteria – syntactic, morphological, and phonological – are clearly missing.

Definiteness may be expressed overtly in Domari through accusative case endings, which distinguish generic or indefinite direct objects from those that are contextually or situationally specified: thus *ama piyami guldas* ‘I am drinking my tea’ (with situational reference to a particular cup of tea), but *ama inmangame’ piyam gulda* ‘I don’t like drinking tea’; *ama šardom plan* ‘I hid the money’, but *šardom ple* ‘I hid some money’. This device is found in other languages that lack overt definite articles but have regular case inflection, such as Hindi or Turkish.

Indefiniteness may be expressed overtly by a postposed indefinite marker *-ak*, which evidently derives from an underlying form of the numeral ‘one’ *\*ek*, reminding of the suffixed indefinite markers of various languages in India such as Sinhalese, Oriya, and Assamese (cf. Masica 1990: 248–250) but also of northern Kurdish (Kurmanji): *dīs-ak kamkaranda* ... ‘one day they were working’, *biddak šar kiy-ak* ‘you want to hide something’, *ama lahedom kažža-k* ‘I saw a man’, *ehra wāšīm qušša-k* ‘something (lit. ‘a story’, < Arabic *qušša*) happened to me’, *fī dēy-ak min dēyeskī l-šīrāq* ‘in one of the towns of Iraq’. Note that the indefinite marker overrides oblique case assignment, which is reserved for definites (*ama lahedom kažža-s* ‘I saw the man’).

Macalister’s (p. 7) example for an indefinite marker<sup>8</sup> – *džur-ik* ‘a woman’ – suggests ambiguity, in both form and distribution, between the

8 When citing Macalister’s data, I adapt the forms to the phonological notation employed here throughout, unless there is evidence for an actual discrepancy in pronunciation; the morphological structure of Macalister’s data is naturally retained.



indefinite marker in *-ak* and the predicative suffixes in *-ik/-ēk* (see discussion below).<sup>9</sup> But consider the following examples:

- (2) a. *hnēn mindird-ik kur-ak*  
 here stand-pred house-indef  
 ‘Here stands a house’
- b. *kur-ik nohr-ik*  
 house-pred red-pred  
 ‘A red house/ the house is red’
- c. *ihī ktīrn-ik*  
 this Christian woman-pred  
 ‘this is a Christian woman’
- d. *erī lāš-ik kišōt-ik*  
 came.f girl-pred small-pred  
 ‘A small girl came’
- e. *pandži nkīs sayyāra-k naw-ik*  
 he at.3sg car-indef new-pred  
 ‘He has a new car’
- f. *ṭōssan bīt-ak, gony-ak qameh, ū bakar-ak*  
 gave.3sg.3pl land-indef sack-indef flour and sheep-indef  
 ‘He gave them a piece of land, one sack of flour, and a sheep’

The predicative suffix is best translated as ‘being’ – *nawik* ‘being new’, etc. – thus attributing a property, while the indefinite marker assigns new-topic status to an entity. Although an analytical separation of the two categories can be maintained, there is nevertheless some potential overlap between them, with indefinite subject nouns in presentative constructions taking either indefinite or predicative markers. It is likely that this overlap in the functional distribution is reinforced by the structural similarity between the two forms. I shall return to the predicative suffixes in section 5.4.

#### 4.2 Nominal inflection

Like other New Indo-Aryan (NIA) languages, Domari shows a layered system of nominal inflection. I use the terminology coined by Masica (1990), where Layer I refers to inflectional elements inherited directly from Old Indo-Aryan (OIA) which in NIA indicate an opposition of nominative and general oblique; Layer II is a closed and limited set of abstract, gram-

9 Littmann (1920: 126) even refers to the predicative endings *-ik/-ēk* as ‘indefinite articles’.

maticalised markers deriving from Middle Indo-Aryan (MIA) postpositions and postposed location adverbs; and Layer III elements are adpositions, i.e. analytical location specifiers.

Domari shares much of its nominal inflection typology with Romani. Both languages have resisted phonological erosion of Layer I markers to a considerable extent. The general oblique, a Layer I marker, serves in both languages not only as the basis for further case formations with Layer II–III markers, as it does in other NIA languages, but also as an independent accusative marker. Gender, like number, is a Layer I property, and is maintained in both languages only in the singular and neutralised in the plural. In both Domari and Romani Layer II affixes are attached directly to, and are inseparable from, the noun, forming in effect a new set of agglutinative synthetic case markers (cf. Friedman 1991 on Romani case). Unlike Romani, however, Domari shows no phonological assimilation to the preceding consonant and so no morphophonological alternation in the forms of Layer II markers (cf. Romani *-kel-ge*, *-tal-da*, etc.). Also in contrast with Romani, the distribution of Layer I–II markers is not constrained by a hierarchy of either animacy or intrinsic referential prominence (see Matras 1997), though referentiality at a local level of the discourse does play a role, with generic nouns being exempted from accusative case marking: *biddī šrīkam mana* ‘I want to buy bread’, vs. *tomis manas* ‘I gave him (my) loaf of bread’. A further significant difference is the fact that Domari allows for the incorporation of possessive personal affixes between the noun stem and Layer II markers (*bāy-im-ke* ‘for my father’). Both Domari and Romani are unique among the NIA languages in having preposed Layer III elements, although only in Romani do they constitute a direct continuation of underlying MIA location adverbs, while in Domari they are entirely borrowed from Arabic. Figure 3 provides an overview of the layout for case formation in Domari:

Figure 3: Case formation

Layer III	Noun stem	Layer I	Layer II
	<i>šōn-</i>	<i>as-</i>	<i>ke</i>
	‘for the boy’ (m.)		
<i>min</i>	<i>kuri-</i>	<i>ya-</i>	<i>kī</i>
	‘from the house’ (f.)		

While on the whole the nominal inflection as described for Domari by Macalister remains intact, a number of differences, some of them significant, can be observed between the corpus considered here and Macalister's discussion.<sup>10</sup> Firstly, some declensional classes are not considered by Macalister, notably feminine nouns in *-a* as well as a number of peripheral classes. Furthermore, a number of genuine Layer II markers appear to have undergone changes in structure and distribution. Finally, attention is given here to the layered structure of possessive markers as well as to a differentiated analysis of the genitive-possessive construction.

#### 4.2.1 Layer I declension classes

Declension groups are distinguished at the level of Layer I markers, as Layer II markers are invariable. An overview is provided in Figure 4:

Figure 4: Declension groups

	sg. nom. - obl./acc.	pl. nom. - obl./acc.
Group 1 'Dom man'	<i>dōm - dōmas</i>	<i>dōme - dōman</i>
Group 2 'person'	<i>manus - mansas</i>	<i>manse - mansan</i>
Group 3 'boy'	<i>šōna - šōnas</i>	<i>šōne - šōnan</i>
Group 4 'girl'	<i>lāši - lāšiya</i>	<i>lāšiye - lāšīyan</i>
Group 5 'Dom woman'	<i>dōmiya - dōmiyē(y)</i>	<i>dōmiye - dōmiyan</i>
Group 6 'city'	<i>ūyar /acc. ūyarī</i>	<i>ūyare - ūyaran</i>
Group 7 'pencil'	<i>qalam - qalamē</i>	<i>qalame - qalaman</i>
Group 8	special cases	

Group 2 differs from Group 1 merely in its syllable structure and the fact that attachment of Layer II markers carries with it a contraction of the noun stem. Groups 5–7 are not considered by Macalister: Group 5 includes feminine nouns in *-a*. We find here apart from ethnicity names in *-iya* (*dōmiya* 'Dom woman', *portkīliya* 'Jewish woman') also Arabic feminine nouns in *-a* (*lamba-lambē* 'lamp', *ḥafla-ḥaflē* 'party'), person names (*zarīfa-zarīfē*) and place names (*yāfa-yafēma* 'in Jaffa', *ḥēfa-ḥēfēma* 'in

- 10 The inherited Layer I vocative, as Macalister (p. 11) remarks, is restricted to words denoting close relationships, where it assumes an ending in *-a* in the masculine sg., and zero-marking for the feminine sg.: *aru xāl-a!* 'come uncle!', *wēštī xālī!* 'sit down aunt!'.

Haifa'). Optionally, the stem may be extended by a carrier consonant: *faransa* 'France', *faransēkī* alongside *faransayēkī* 'from France'.

Group 6 is characterised by consonantal stems that have strictly speaking no distinct Layer I marker in the singular: *ūyarma* 'in the city', *ḥaramta* 'to the Mosque compound'. However, they appear to have a highlighted accusative form in *-ī*: *ama lahami ūyarī* 'I see the city', *qol šibbākī!* 'open the window!', *lahedom ḥṣānī* 'I saw the horse'. The noun *džuwir* 'woman' partly belongs to this group too; it appears in the accusative as *džuwirī*, though the presence of a Layer I oblique marker as a base for Layer II elements is variable: *džuwirkī* alongside a shift to the *-i*-class in *džuwrakī* 'from the woman'. The exact status and origin of this *-ī* are not sufficiently clear at this stage, though I speculate that we are dealing here not with a genuine Layer II case marker, but with an attempt to imitate the extended structure of the accusative direct object of other declension classes. This is achieved through attachment of the predicative suffix *-ī* that follows consonantal noun stems, and which generally appears in presentative constructions, thus: *ama piyami ḥalībī* 'I drink [this is] milk'.<sup>11</sup>

Group 7 seems to include primarily loan elements and proper nouns (*?āmālēke* 'for Amāl'). Special cases of nominal declension include *zara-zares*, pl. *zare-zarten* 'boy', already mentioned by Macalister. Another particular case is *dōmarī-dōmarīyasma* 'in Domari'. Noteworthy is also the insertion of an accented vowel – often reduplicating a final stem vowel – as an oblique base marker for Layer II elements with place names ending in a consonant: *ṣammānāma* 'in Amman', *ṣīrāqāta* 'to Iraq', *min qudsāki* 'from Jerusalem (< Arabic *al-quds*)', *liddīta* 'to Lidd'.

#### 4.2.2 Forms and functions of Layer II markers

Domari has, in principle, five distinct Layer II markers, for which Macalister uses the following terms (forms in brackets as cited by Macalister): Dative (*-ta*), Locative (*-ma*), Directive (*-kera* or *-ke*), Associative (*-sanni* or *-san*), and Ablative (*-k* or *-ki*). Missing from the inventory is a marker that would correspond to the Romani or Hindi adjectival Genitive, a category unknown to the Domari system. The forms themselves, and to some extent their distribution as found in our corpus differ partly from Macalister's description. As for the terminology, for the sake of both adequacy and consistency I continue to use the labels introduced by

11 Macalister (p. 9) has interpreted this as an ending attached to Arabic loan nouns; cf. discussion below.

Macalister, with the exception of the Directive, for which I find the term Benefactive more appropriate.

The Dative in *-ta* generally expresses contact which does not explicitly entail containment. With location expressions and verbs of motion it expresses the goal of a motion – *garom kamasta* ‘I went to work’, *biddī džam kuryata* ‘I want to go home’, *tirdom kubayē ṭawlēṭa* ‘I put the cup on the table’ – or the location of a state: *lakedomsi wēsrēk kursata* ‘I have seen him sitting on the chair’. Further types of contact expressed by the Dative can be with instruments – *mamnūṣī xūlšad goryanta* ‘they are not allowed to ride horses’, *ṣazīfkandi rabbabēta* ‘they play the rabbab’ – or among humans: *ṭarrafrēn baṣḍēmanta* ‘we met one another’. Finally, the Dative can also express an abstraction analogous to actual contact: *smari dōmanta* ‘he hears about the Dom’, *širdom abuske putrēmta* ‘I told him about my sons’, *dawwirkaradi putrosta* ‘they are looking for his son’, *laḡiškade ... ehe raqqāṣanta* ‘they had an argument ... about those dancers’. The Dative in *-ta* is confined to a group of speakers, generally the oldest among the fluent speakers, while in the speech of the others this form has been entirely replaced by *-ka*, which covers exactly the same functions. This development appears to be the result of a levelling within the Layer II paradigm, triggered by the presence of two other forms in *-k-*, namely the Benefactive (Macalister’s ‘Directive’) in *-ke(ra)* and the Ablative/Prepositional in *-kī*.

The Locative in *-ma*, by contrast, expresses contained location, either stative – *šaryandi kuryisma dōmankī* ‘they are hiding in the houses of the Doms’ – or directional – *ere hindar ūyarma* ‘they came here into the town’. Here too, analogous abstractions can be found: *kayma kallamōk atu? dōmasma!* ‘What are you speaking (in)? In Domari!’. The so-called Directive is found to have a strictly Benefactive function, with no intrinsic physical movement associated with it: *širdom dāyimke* ‘I said to my mother’, *tu qayiš putrimke!* ‘serve food for my son!’, *tomis ple ṣadīqimke* ‘I gave money to my friend’. Like the Romani dative, which is its cognate, it is used in conjunction with just a small, closed set of verbs. The long form cited by Macalister – *-kera* – has virtually disappeared and is documented only in the pronominal form *amakera* ‘for me’, and even there the more frequently used is the contracted form *amake*. The Associative is reported already by Macalister to be variable in both distribution and form, being gradually replaced by a preposition (Arabic *maʿ* ‘with’) and assuming a reduced form in *-san*. In our corpus only few instances of the Asso-

ciative could be found, and only in the reduced form: *širdom bāyimsan* ‘I spoke with my father’.

Finally, the Ablative is found only in the longer form reported by Macalister, namely *-kī*. As an independent marker of semantic case expressing source, it is found only among the older speakers, and, it seems, only in expressions implying initial containment – *kildom kuryakī* ‘I went out of the house’ – while non-containment is expressed through an added preposition (Arabic *min* ‘from’): *sindom min zareskī* ‘I heard from the boy’. Here, the preposition must not be interpreted as merely reinforcing the synthetic Ablative marker. Rather, the Ablative serves as a Prepositional case: *ama xarrifrom maʕ šāḥbimkī* ‘I spoke with my friend’ (cf. Associative *šāḥbimsan*), *ama garom la kuryiskī* ‘I went to his house’ (cf. Dative *kuryista*). For the younger among the fluent speakers, who have generalised the use of the ablative preposition *min*, the Ablative no longer has an independent semantic function and is confined to this use as a Prepositional case.

#### 4.2.3 Possessive inflection

Macalister downplays somewhat the regularities of possessive inflection in Domari, speaking of an ‘all but completely arbitrary’ pattern of vocalisation preceding and following the consonantal suffixes that serve as possessive markers (p. 21). This confusion is due in part to actual systematic differences among the different classes of pronominal markers in their different roles (possessors, direct objects, prepositional objects), and partly to Macalister’s ignorance of the tense system and external tense affixes, which may attach to object clitics, thus overriding epenthetic vowels that otherwise connect them to the subject concord markers (*lahedom-is* ‘I saw him’, *lahedom-s-i* ‘I have seen him’, *lahedom-s-a* ‘I had seen him; cf. discussion below). Finally, Macalister appears to ignore the layered structure of possessive affixes themselves.

In general, the consonantal person markers in Domari show consistent and symmetrical forms: 1sg *-m*, 2sg *-r*, 3sg *-s*, 1pl *-man*, 2pl *-ran*, 3pl *-san*. Possessive markers are based on these consonantal forms, which are added to an attaching vowel, a first-layer possessive marker as it were. At the level of this first layer, number and case of the possessed noun are distinguished. Number distinction only appears in the nominative (Macalister does note the number distinction, for which he cites similar data): *kuryos* ‘his house’, *kuryēs* ‘his houses’; *putrom* ‘my son’, *putrēm* ‘my sons’; *dīrom* ‘my daughter’, *dīrēm* ‘my daughters’. Case distinction

involves oblique marking in the first (or attachment) layer for possessive markers when the possessed noun appears in non-subject position: *bāyom* ‘my father’, but *lahedom bāyim* ‘I saw my father’; *kuryom* ‘my house’ and *kuryor* ‘your house’, but *garom min kuryimkī la kuryirkī* ‘I went from my house to your house’. The origin of this possessive case inflection is not clear, but it could derive from some form of relativiser or determiner which once mediated between the head noun and a postposed possessive pronoun, agreeing with the head in number and case. Gender agreement may have been levelled at a later stage. The erosion and simplification of this paradigm is still ongoing, and we only find case distinctions in the singular forms – *-om*, *-or*, *-os* vs. *-im*, *-ir*, *-is*, – while the plural forms are, so far, only documented with a single vowel attachment (*-oman*, *-oran*, *-osan*) for nouns in different thematic roles.

#### 4.3 The genitive-possessive construction

The Domari genitive-possessive construction is based on a generalisation of the 3sg possessive marker. It employs the singular possessive marker on the head, irrespective of the actual number of the possessor-determiner, while the determiner itself appears in the ablative-prepositional case (possibly replacing an underlying genitive case in similar function and form). The word order in this format is consistently head-determiner.

- (3) a. *kury-os*      *kažž-as-kī*  
          house-poss man-obl.m-abl  
          ‘the man’s house’  
   b. *grawar-os*   *dōm-an-kī*  
      chief-poss   dom-obl.pl-abl  
      ‘the leader (chief, or Mukhtar) of the Doms’

A comparison with genetically related, as well as with contiguous languages, namely Hindi and Romani, Kurdish, Arabic, and Turkish, which have quite distinct types of genitive constructions, shows that Domari has a rather unique structure which is typologically most akin to the emerging and still very much marked Arabic construction *bēt-o li-z-zalame* ‘the man’s house’, lit. ‘his house of the man’, where the head carries cataphoric reference to the determiner through the possessive suffix that agrees with it, and the determiner follows the head and is marked for its role as possessor by an indirect case marker. The order head-det is also shared with Kurdish and other Iranian languages, while on the other hand phoric reference on the head (through a possessive affix) to the determiner-possessor is found

in Turkish. None of the features of the Hindi and Romani construction, which show adjectival agreement on the preposed determiner with the postposed head (Hindi *laṛk-ō-kā ghar*, Romani *le rakl-en-go kher* 'the boys' house'), can now be found in Domari, with the exception of the universal marking of the determiner-possessor through some form of an oblique case. The determiner-head order on which Macalister (p. 13) reports (*kuryak kapyos* 'the door of the house'), and which agrees with Romani and Hindi, could not be encountered in the corpus. If indeed a shift in word order has recently taken place, then it is likely to be a result of convergence with Arabic.

If the determiner itself is marked for possession in a multiple possessive construction, then the possessive affix may, variably however, carry the oblique form: *bāy-os ṣādīq-im-kī* 'my friend's father', but also *kury-os bār-om-kī* 'my brother's house'. If the head is not in subject position, it takes whatever case reflects its syntactic role; a non-nominative case will then trigger an oblique form of the possessive marker on the head: *kury-os* 'his house', but *ama tirdomi kury-is-ma bār-om-kī* 'I live [= have settled in] my brother's house'; *zaman-is-ma nohr-an-kī* 'in the time of the British (rule)'. Note that Layer II case affixes, as mentioned in the introductory remarks to this section, follow the possessive marker (*kury-is-ma* 'in his house'). This differs of course from Romani, which has no clitic person affixes, and which generally does not allow any insertion between Layer I and Layer II markers, while on the other hand Domari is stricter than Hindi in that it only allows person affixes, but not any other determiners to intervene between the head and Layer II affixes. In a sense, then, the evolution of Layer II items as synthetic markers in Domari is at an intermediate stage compared to the two other languages: They are fully grammaticalised and cannot 'float' within the noun phrase, while on the other hand they show at least two types of distributions, attaching to either Layer I oblique nominal affixes in the strict sense, or to the oblique forms of the possessive person clitic.

A final remark on the genitive-possessive construction concerns what Macalister (p. 14) attempted to identify as a Persian or Persian-type Izafe construction in *-i*. According to Macalister, this may take on two distinct forms: (i) The plain Izafe consists merely of an insertion of the Izafe marker in-between the head and the determiner-possessor: *siri-i-manus* 'the man's head'. (ii) The second type is a contamination of the Izafe with the inflected genitive-possessive construction outlined above, namely



*siryos-i-manusask*. As far as the first structure is concerned, I have been unable to find any trace of it whatsoever, and it is not clear whether it has since perished, or whether some kind of misinterpretation might be involved. As for the ‘contaminated’ structure, the data appear to be quite straightforward in suggesting that this has, in fact, little to do with a Persian (or rather, Iranian; Persian itself has *-e*) Izafe structure. Rather, we are dealing, once again, with the predicative suffix attached to consonantal stems, namely *-ī/-i*. Consider the following examples:

- (4) a. *īhī kuryom-i*  
           this house.1sg.poss-pred  
           ‘This is my house’  
       b. *īhī kuryos-i*                    *bāyimkī*  
           this house.3.poss-pred father.1sg.poss.abl  
           ‘This is my father’s house’  
       c. *šurūsos*                    *dōmankī šādī-k*  
           wedding.3.poss Dom.pl.abl normal-pred  
           ‘Dom weddings are conventional’  
       d. ... *dfāškar adžaros kuryakī*  
           pay.subj rent.3.poss house.abl  
           ‘... in order to pay the house rent’

In presentative constructions, a predicative suffix is normally attached to the head of the predication. In (4a) the head is a plain noun inflected for a 1sg personal possessor. In (4b) it is the head of a genitive-possessive construction, precisely the type suspected by Macalister to be a Persian Izafe contamination. Note that all presentative constructions in the corpus that involve a head that is also the head of a genitive-possession construction, i.e. of the type illustrated in (4b), show such attachment of a predication marker. The choice of *-ī/-i* for a predication suffix is conditioned by the consonant ending of the 3rd person possessor. The structure is thus analysable and predictable. Note finally that in cases where the genitive-possessive construction does not constitute the pivot of a presentative predication, as is the case in (4c–d), no predicative suffix is attached to it.

Alongside the principal genitive-possessive construction, a morphologically ‘weaker’ form expressing multiple possession can be found. It involves a determiner-possessor that is inflected for person, preceding a head that lacks phoric reference to the possessor: *bāyim kuri* ‘my father’s house’. From a comparison with related and contiguous languages, it

would seem that this might represent a simplified form of an underlying det-head construction inherited from Indic; its distribution in the corpus however does not quite support such an interpretation, as the construction seems to surface more frequently among less-fluent speakers. The analytic genitive in *kāk-*, cited by Macalister, appears sporadically – *tomis giš plēm kākīm* ‘I gave him all my money’. Noteworthy is that, although at first glance this seems to copy the Arabic analytical genitive-possessive in *tabaʕ-*, albeit based on an indigenous particle most likely of deictic-relative origin, the possessive inflection on *pl-ēm* ‘my money’ is nevertheless retained. It is yet to be established whether this has contrastive function (as in Arabic, *bēt-ī tabaʕ-ī* ‘my own [nobody else’s] house’, cf. *bēt-ī* or *l-bēt tabaʕ-ī* ‘my house’).

#### 4.4 Location expressions

Macalister does not devote any particular section to this area, and so it deserves some remarks. Firstly, Domari employs Arabic prepositions which constitute the only Layer III-type modification to the noun. Some Arabic prepositions still compete with Layer II case affixes. Thus we find *maʕ* ‘with’, *min* ‘from’, *la* ‘to’, *fī* ‘in’ and *ʕind* ‘at’ competing with the Associative, (independent) Ablative, Dative, and Locative respectively. Occasional doubling may be observed (*fī ʕareʕma* ‘on the street’), though on the whole prepositions trigger the use of the Ablative as a Prepositional case. We have a parallel in Romani, where the Locative case in *-tal-da* has been generalised in most dialects as a prepositional case. Noteworthy however is that while Romani recruits Layer III elements largely through grammaticalisation of inherited stock items, in Domari Layer III has undergone complete fusion with the counterpart class of elements in Arabic.

Second, mention must be made of genitive-possessive location expressions. They consist of a modifier location adverb inflected for (oblique) possession and Locative case, preceding a head in the Ablative/Locative: *mandžisma kuryakī* ‘inside the house’ (lit. ‘in its-inside from-the-house’). The general format could well be inherited from Indo-Aryan (cf. Hindi genitive-location constructions, where however the location adverb itself does not appear in a possessive form). A closer structural parallel can be found in Turkish, where however word order differs, the head preceding the modifier. Arabic location expressions may substitute for inherited ones, but they are still inserted into the same type of construction: thus *šanšisma kuryimkī* ‘next to my house’, alongside *žambisma lāčaki* ‘next to the girl’ (< Arabic *žamb* ‘next to’).

Third, a simplified form of such relational constructions is found, where the head is marked for the Locative and is followed by an uninflected location adverb: *kuryama mandža* 'inside the house', *kuryama bara* 'outside the house'. There is only a very small set of such inherited location adverbs, all expressing strict spatial relations: *mandža* 'in', *bara* 'out', *paš* 'behind', *agir* 'in front', *atun* 'above', and *axār* 'below'. Temporal and more specified spatial relations are expressed through Arabic items.

Finally, there is a class of local relation expressions that only appear with person markers, functioning so to speak as indirect object pronouns, and deriving from underlying prepositions that have only survived in person-inflected form. Most of these have been noted by Macalister (p. 20–21) in his discussion of the pronoun paradigm, though some differences are apparent. There are in principle three strategies for the formation of such expressions. The first involves the plain attachment of a person affix from the set of oblique clitic pronouns to a local expression: *wāšī-m* 'with me', *wāšī-r*, 'with you'; *nkī-s* 'by him/in his possession', *nkī-man* 'by us', and so forth. The second involves the attachment of Layer II case markers in conjunction with local expressions, notably for the Benefactive in *ab-us-ke* 'for him', *ab-ran-ke* 'for you (pl.)', and in Macalister's data also for the dative *at-ur-ta* 'for you', *at-san-ta* 'for them'. The final strategy is restricted to first person markers and involves simply the addition of a Layer II case suffix to the independent pronominal form: *ama-ke* 'for me' (Macalister: *ama-kerā*), and in Macalister's data also *ama-ta* 'to me' and *ama-ma* 'by me/in my possession'.

The following differences can be noted, compared to Macalister's material: The Dative form for the 2nd and 3rd persons in *at-ur-ta*, *at-us-ta* etc. is replaced by the form *atnī-* (cognate with *atun* 'above') with no Layer II additions (*atnī-s* 'about him', *atnī-r* 'about you'), while the first person forms *ama-ta*, *emin-ta* remain part of the paradigm. Semantically, the Dative is restricted to abstract contact. Actual physical contact is always expressed by the Locative, which now has forms in *nkī-* throughout (*nkī-m* 'to me/at my disposal'), with no trace of the separate 1sg *amama* cited by Macalister. The Associative shows the same forms as in Macalister's description. The Benefactive forms are always reduced, but continue to show separate types for the 1st person (*amake*, *eminke* 'for me, us', but *abuske*, *aburke*, *abranke*, *absanke*). Macalister's Ablative forms in *mneš-* have been replaced by either *minšī-* or *mēšī-*. Finally, the Arabic preposition *ʕan* 'about' has been integrated into the paradigm, taking on the form

*ʕankī-* (*ʕankīm* ‘about me’ etc.), which competes with the other Dative forms.

#### 4.5 Pronominals and adnominal agreement

The independent, nominative pronouns for the 1st and 2nd persons (*ama*, *atu*, pl. *eme*, *ime*) are straightforward derivations of Indic pronouns. For the 3rd person, Domari has *pandži* (sg.) and *pandžan* (pl.), which are derived from reflexives in an etymological oblique form *appñ-*. The identical form is documented as a 3rd person reflexive for Transcaucasian Karači by Patkanoff (1907/1908: 262): *Hye duhend baŋgi khasta* ‘they wash[ed] their hands’. The renewal of the 3rd person pronoun allows us to draw an isogloss, based on published wordlists, between the northern Domari-type varieties, which employ the remote demonstratives in *hu/hi*, and the southern varieties in Lebanon – cf. Groome’s (1891: 25) list of items recorded in Beirut – and Palestine, which have *pandži/pandžan*.

The demonstratives in *h-* are nevertheless present in Jerusalem Domari. Macalister (p. 23) mentions an opposition of distance – between *aha* and *uhu* – though only for the masculine singular. There is indeed only one nominative form for the feminine and plural respectively. However, in the oblique forms, which Macalister does not discuss, the distance opposition re-surfaces (Figure 5):

Figure 5: Demonstratives

	Proximate nom. - obl.	Distal nom. - obl.
m.sg.	<i>aha - ēras</i>	<i>uhu - oras</i>
f.sg.	<i>īhī - ēra</i>	<i>īhī - ora</i>
pl.	<i>ehe - ēran</i>	<i>ehe - oran</i>

Only *aha/ēras*, the proximate m.sg, appears to be in use in a pronominal function. The full paradigm has adnominal (attributive) function and its forms display the same agreement patterns as adjectives: *tilla zara* ‘the big boy’, *tillī lāšī* ‘the big girl’, *tille zare* ‘the big children’. As in Romani, gender agreement is neutralised in the plural.<sup>12</sup> Unlike Romani, Domari adjectives and demonstratives do not show case agreement.

12 I thank Viktor Elšík (personal communication) for pointing out that this feature appears to be restricted to Romani and Domari among NIA languages.

5. VERB MORPHOLOGY

5.1 Categories and linear arrangement

The Domari verb is characterised by its retention of MIA person affixes in the present conjugation, and the emergence of a perfective conjugation, through the attachment of person affixes to the historical past participle in *-ta*. Both phenomena are shared with Romani, although the actual forms differ in part.

The verb stem occupies the first or left-most position in the linear blueprint for the Domari verb (see Figure 6). It may be followed by derivational extensions expressing transitivity (i.e. causative, usually in *-naw-*) or de-transitivity (i.e. passive, in present *-y-*, past & subjunctive *-ī-*). This derivation, claimed by Macalister (p. 31) to be rare, is found in our corpus to be quite productive: *ban-ari* ‘he shuts’ > *ban-y-ari* ‘it is being shut’; *šar-dom* ‘I hid (tr.)’ > *šar-ī-rom* ‘I hid (intr.)’; *qē-ror* ‘you ate’ > *q-naw-idor* ‘you fed’, etc.

Aspect consists of the opposition between progressive (or non-completion), expressed by the present, imperfect, subjunctive; and perfectivity (or completion), expressed by forms based on the historical past participle – preterite or ‘unspecified perfective’, perfect, pluperfect. Perfective categories are formed through an extension to the verb stem in *-d-* or *-r-*, derived from MIA *-t-*.

‘Mood’ refers here (Figure 6) to the explicit marking of the subjunctive/optative. This is only applicable to some verbs which employ an historical optative extension in *-š-* for this purpose. For other verbs, the subjunctive is identical to the present indicative, except for its lack of tense specification. Other verb classes have generalised the use of the historical optative extension in *-š-* to indicative forms as well (see below).

Figure 6: Linear outline of the verb

stem	derivation	aspect/mood	subject	object	tense
<i>xiz -</i>	<i>naw-</i>	<i>(i)d-</i>	<i>om-</i>	<i>san-</i>	<i>a</i>
‘I had made them laugh’					
<i>lah-</i>			<i>ad-</i>	<i>man-</i>	<i>i</i>
‘they are seeing us’					
<i>bag-</i>	<i>y-</i>		<i>ar-</i>		<i>a</i>
‘it used to break’					
<i>šar-</i>	<i>ī-</i>	<i>š-</i>	<i>am</i>		
‘(that) I hide’ (intrans.)					

There are two sets of subject concord markers. The progressive categories take 1sg *-m*, 2sg *-k/-ī/-ø*, 3sg *-r*, 1pl *-n*, 2pl *-s*, 3pl *-(n)d*; note the diversity of 2sg markers, which vary according to conjugation group. This set is in principle a continuation of MIA concord markers (but see Turner 1926: 280). The perfective set is *-om*, *-or*, *-a/-ī*, *-ēn*, *-ēs*, *-e*. Here the 3rd person markers are adjectival, reflecting the historical past participle which has become an active finite form. Unlike agreement in the ergative NIA languages, agreement in Domari is with the subject in both transitive and intransitive verbs: *gara* ‘he went’, *garī* ‘she went’; *laherda* ‘he saw’, *laherdī* ‘she saw’.

The subject concord marker in the 3sg of the perfective assumes the form *-os*, identical with the possessive marker, when an object pronominal clitic is present (*laherdos-im* ‘s/he saw me’). Barr (1943) attributes this to Iranian influence triggering different agreement patterns with transitive and intransitive verbs, and featuring object pronoun markers in subject position with transitive predicates. Whether this is the only motivation for the split in 3sg concord markers, is questionable. In the 3pl, a reduplication of the perfective extension may appear before object clitics: thus *laherde* ‘they saw’, but *laherde-d-is* ‘they saw him/her’. Here the insertion is clearly phonologically motivated. Nevertheless, the appearance of potentially competing forms in the 3sg can be taken as a sign of the gradual retreat of the active participle of the Hindi type. In Romani, active participles with adjectival agreement compete with person-inflected forms in *-as*. Adjectival agreement is restricted to unaccusative verbs in the southern Balkan dialects of Romani (*dikhlas* ‘he saw’ but *avilo* ‘he came’); in Transylvanian dialects it expresses a kind of evidentiality (*avilas* ‘he came’, *avilo* ‘he came suddenly/unexpectedly’; see Matras 1995), while in the central and northern European dialects of Romani, active 3sg adjectival agreement cannot be found at all.

Object person markers consist of the same set of affixes used with prepositions and in possessive constructions, albeit reduced to consonant-initial forms. A syllable boundary is created either through the attachment of a tense affix – *laham-r-i* ‘I see you’, *laherdom-r-i* ‘I have seen you’ – or, in the absence of such a tense marker, namely in the subjunctive and the preterite (unspecified perfective), by an epenthetic vowel *-i*: *ta laham-ir* ‘so that I may see you’.

The final, right-most position in the verb layout is occupied by tense affixes. There are two such affixes, which I call the contextualising marker

(-i) and the de-contextualising or remoteness marker (-a). The contextualising marker figures in the present (*laham-i* ‘I see’) and perfect (*lahedom-i* ‘I have seen’). Its function is the actualisation of an action or its result within the currently activated context of the speech event. The de-contextualising marker forms the imperfect when added to the present form (*laham-a* ‘I was in the habit of seeing’), and the pluperfect when added to the unspecified perfective form (*lahedom-a* ‘I had seen’). Its function is to emphasise the demarcation between the action conveyed by the verb, and the currently activated speech context. Note that tense differs from the aspectual qualities of progressivity (present as well as imperfect) or perfectivity (plain perfective as well as pluperfect).

Figure 7: Overview of TAM categories – *lah-* ‘to see’

Subjunctive	Present	Imperfect
<i>laham</i>	<i>laham-i</i>	<i>laham-a</i>
Unspecified Perfective	Perfect	Pluperfect/Counterfactual
<i>lahedom</i>	<i>lahedom-i</i>	<i>lahedom-a</i>

Two categories are unmarked for tense: The subjunctive does not refer to the speech event but shows syntactic dependency on a modal expression or main clause verb, sharing a non-resultative or non-completed reading with the present and the imperfect. The simple past, or preterite, conveys the completion of an action in absolute terms, without specific reference to the speech context; it therefore seems appropriate to adopt Masica’s (1990) term ‘unspecified perfective’ here, especially owing to the parallels with other NIA languages.

The present and subjunctive are also used for future time reference. No trace could be found in the corpus of the future which Macalister (p. 28) describes as having an insertion in -y-; it seems that Macalister had mistaken the de-transitivised (= passive) form in the present tense, in an instance where it carried future time reference, for a morphological future (admitting nonetheless that ‘the use of this form is not very common, and as a rule the sense of futurity, if desired, is left to the hearer to infer from the context’). This impression is reinforced by Macalister’s failure to recognise the present passive formation in -y- (*nanami* ‘I bring’ > *nanyami* ‘I am being brought’), and his mistaken attempt instead to reconstruct a non-existent present passive, deriving it from the past passive (\**nanīrami* ‘I am being brought’, from *nanīrom* ‘I was brought’; p. 32).

As a final remark on the verbal layout one should note the grammaticalised double negation in the present tense: *mangamsani* 'I like them' > *(i)nmangamsane* 'I don't like them'. Its initial component contains the inherited negation marker *n-*, usually preceded by an epenthetic vowel *i-*. Its final component is an accented *-e'* ending in a glottal stop, a rare phoneme in the pre-Arabic component; this latter morpheme, of unknown origin, appears to have been a reinforcer negative marker which has now assumed the role of a principal negator (compare French *pas*, and Arabic *-iṣ*). In simplified negative constructions, only the second component appears: *piyame* 'I don't drink'. The negated form of *ašte* 'there is' is *nhe*'. Other tenses generally take the negator *na*, the imperative can take *ma*. Domari also employs Arabic negators. Arabic *mā ... -iṣ* or either one of its two components attaches to the inflected Arabic verbs *kān-* and *ṣār-* and to the quasi-verb *bidd-*. Non-verbal predications are negated using Arabic *miṣṣ*.

## 5.2 Inflection groups

Macalister's treatment of verb inflection groups is, like his discussion of tenses, rather superficial. He merely distinguishes two types of preterite formations, in *d-* and in *r-* (p. 29), and goes on to name eight irregular verbs, mentioning in passing 'many little irregularities ... which are hardly of sufficient importance to enumerate'. The Domari verb inflection system constitutes in actual fact a rather entangled web of patterns. Among the synchronic parameters that condition affiliation to an inflection type are transitivity/intransitivity (essentially a reflection of historical phonological processes, however), the presence of an intensifier perfective marker, variation in the existential verb, syllable reduction and epenthesis, shifts between the historical optative and indicative paradigms, suppletion, and quasi-suppletion.



Figure 8: Principal verb inflection groups

	Group 1 transitives (cons. stem)	Group 2 transitives with perf. intensifier	Group 3 de-transitives (-ī- extension)	Group 4 existential derivations
	<i>ban-</i> ‘shut’	<i>dow-</i> ‘wash’	<i>šar-y-</i> ‘hide’	<i>skunn-</i> ‘dwell’
1sg present	<i>ban-ami</i>	<i>dow-ami</i>	<i>šar-y-ami</i>	<i>skunn-(h)omi</i>
1sg subjunctive	<i>ban-am</i>	<i>dow-am</i>	<i>šar-ī-šam</i>	<i>skunn-hōšam</i>
2sg present	<i>ban-ēk</i>	<i>dow-ēk</i>	<i>šar-y-āk</i>	<i>skunn-(h)ok</i>
2sg imperative	<i>ban!</i>	<i>dow!</i>	<i>šar-ī-ši!</i>	<i>skunn-hōši</i>
1sg perfective	<i>ban-dom</i>	<i>dow-ir-dom</i>	<i>šar-ī-rom</i>	<i>skunn-(ah)rom</i>

My point of departure is a division into four principal inflection groups (Figure 8). Group 1 consists of transitive verbs whose stems end in consonants. It is characterised by the use of a syncopated form for the subjunctive, the ending *-ēk* in the 2sg present, the bare stem form in the 2sg imperative and subjunctive, and a perfective extension in *-d-*. Subgroups of Group 1 include (i) stems ending in a velar or glottal consonant, where an epenthetic vowel appears in the perfective (*bag-ami* ‘I break’, *bag-i-dom* ‘I broke’; *lah-ami* ‘I see, *lah-e-dom* ‘I saw’); to these belong loan verbs and secondary verbal derivations in *-k-*, a reduced form of *kar-* ‘to do’, as in *lagiš-k-ami* ‘I fight’, *lagiš-k-idom* ‘I fought’; *š(t)rī-k-ami* ‘I buy’, *š(t)rī-k-idom* ‘I bought’; (ii) stems ending in nasals, which show syllable reduction in the present tense (*sn-ami* ‘I hear’, *sin-dom* ‘I heard’; *kn-ami* ‘I sell’, *kun-dom* ‘I sold’); and (iii) causatives in *-naw-* etc. (also *-law-*, *-raw-*): *q-naw-ami* ‘I feed’, *q-naw-idom* ‘I fed’. Group 2 is essentially identical with Group 1 except for the presence of a perfective intensifier *-ir-* preceding the perfective extension.

Group 3, characterised by a perfective extension in *-r-*, consists of de-transitive verbs formed by means of an extension *-y-* to the stem in the present, and *-ī-* in the perfective and subjunctive. It is apparently this vowel ending of the stem that triggered the shift from the historical participle ending in *-ta* to *-ra*, contrasting with *-da* in consonantal stems (cf. *pī-rom* ‘I drank’, *ga-rom* ‘I went’, but *kar-dom* ‘I did’, *sin-dom* ‘I heard’; cf. also Romani *pi-lom* ‘I drank’, *ge-lom* ‘I went’ etc., but *ker-dom* ‘I did’, *šun-dom* ‘I heard’). Further features of Group 3 are the subjunctive/imperative

extension in *-š-*, likely to be derived from an underlying auxiliary in *(a)ččh-* ‘to stay’, the 2sg present in *-āk*, and the 2sg subjunctive/imperative in *-ī*.

Group 4 finally consists of derivations of the existential verbs, which include the enclitic copula, the independent existential verb ‘to become’, and numerous Arabic loan verbs that employ an existential verb as a carrier verb (see discussion below). Here, progressive forms take the vowel *o/ō* of the existential verb (< *ho*) following the stem, the subjunctive/imperative takes *-š-*, and the perfective extension is *-r-* (as with other vocalic stems).

A fifth class can be defined as including assorted particular, irregular, or isolated cases: (i) The verb *piyam/pīrom* ‘I drink/I drank’ assumes an intermediate position between Groups 1 and 3; it is transitive, has a 2sg present in *-ēk*, a plain subjunctive (*biddī piyam* ‘I want to drink’) and imperative (*pī!* ‘drink!’), but a perfective extension in *-r-*. This is due to the phonological parallels with Groups 3 and 4, namely the presence of a vowel stem, which however is unusual in a transitive verb. (ii) A series of verbs, typically expressing motion or state, show infiltration of the underlying optative/subjunctive in *-š-* (from < *č* < *aččh-*) into the indicative paradigms, thus *nikšami* ‘I enter’, *šūšami* ‘I sleep’. (iii) Historical phonological developments result in stem alternations between the present/progressive and past/perfective paradigms (quasi suppletion): *qumn-/qēr-* ‘to eat’, *šūš-/sit-* ‘to sleep’, *nast-/nasr-* ‘to escape’, *xašt-xazr-* ‘to laugh’, *nik(š)-/nigr-* ‘to enter’, *dē-/t-* ‘to give’. (iv) Genuine suppletion is encountered with *aw-/ēr-* ‘to come’, *dža-/gar-* ‘to go’. (v) Finally, we have isolated cases in *sakami/sakarom* ‘I can/could’, and *kahíndomi/kahíndirom* ‘I look/looked’, the latter unique also in its accentuation.

### 5.3 Functions and distribution of tenses

The external tense markers *-i* (contextualising) and *-a* (de-contextualising, remoteness) are ignored by Macalister, who merely notes (p. 29) that an *-i* ending may occur which has ‘no traceable difference in sense or use from the form without *i*’.<sup>13</sup> Macalister’s tense paradigm thus includes merely the present and the preterite (unspecified perfective).

The most challenging opposition among the tense forms, as far as function and distribution are concerned, appears to be that between the unspecified perfective (*kardom* ‘I did’) and the perfect (*kardomi* ‘I have

13 Pott (1846: 182) however had noted the use of *-a* as an ‘imperfect’ marker on the copula in Syrian Domari (*stēmi-stēma* ‘I am-was’), but did not comment on lexical verbs.

done'). This is due to their interchangeability in a variety of contexts. The perfect however clearly emphasises the immediate contextual relevance of an accomplished event, state, or action. This may result in a lexical-semantic differentiation conveyed by the opposition of tense forms, as in *džanami* 'I know', *džandom* 'I knew', *džandomi* 'I have (now) understood'. With verbs expressing state and condition, the perfect usually represents actuality of the accomplished state; consider *weštami* 'I sit (habitually)', *wēsromi* 'I am seated'; *ama kamkame' yašnī wēsromi* 'I do not work, that is, I am retired' (lit. 'I have sat down'); *bag-* 'to break' > *bagīromi!* 'I am exhausted' (lit. 'I have been broken').

In questionnaire elicitation, the perfect is generally chosen by speakers to translate Arabic present participles expressing state: *inte nāyem* 'you are asleep' > *atu sitori*, while the present *šūšēk* 'you sleep' has habitual meaning, and the unspecified perfective *sitor* may refer to a state that is not contextually relevant, that is, does not extend into the present speech situation. For verbs indicating specifically a change of state, there is a strong tendency for the perfect to take over all instances of immediate contextual relevance, reducing the present to habitual readings; thus *da'iman byami* 'I am always frightened', but *heššaš bīromi* 'I am now frightened'. Most notably, this tendency appears in the existential verb, where the present paradigm *homi*, *hok* 'I am, you are' etc. has been almost entirely replaced by the perfect of the verb 'to become' – *(a)hromi*, *(a)hrori* etc. – in enclitic position (see discussion below).

Rather straightforward is the use of the remoteness marker in *-a* with both existential and lexical verbs. In the existential verb, the underlying perfect, now functioning as a present, changes into a past copula when a de-contextualising suffix is added – *(a)hroma* 'I was'. It is often supplemented by an inflected form of the Arabic past-tense existential verb *kān*: *lamma kunt kaštutahroma* ... 'when I was young'. With lexical verbs, the imperfect has strictly a habitual meaning: *bāyōs kamkara baladiyyēma* 'his father used to work for the municipality'. Here too, the Arabic copula is often used to reinforce the imperfect: *kull dīs kunt džama ūyarta* 'I used to go to town every day'. The pluperfect, which also relies on the remoteness or de-contextualising suffix *-a*, is most frequently encountered in counterfactual constructions. Here, an uninflected form of Arabic *kān* is used, copying the use of uninflected *kān* in such constructions in Arabic: *law wāšim iple kān tomra* 'if I had had money I would have given it to you'

(Ar. ...*kān aṣṭētak*); *law ērom xužoti kān laherdomsa* ‘if I had come yesterday I would have seen him’.

#### 5.4 Existential predications

Reference has already been made above to the enclitic impersonal, non-verbal predication marker in *-ēk/-ik* (sg), *-ēni* (pl), discussed by Macalister under the heading ‘predicative suffix’. These forms resemble the indefinite article in having an ending *-k*, and indeed Littmann (1920) implies, based on his reading of Macalister, that the two categories are interchangeable. The set in *ēk/ik/ēni* actually only attaches to vowels. Contrary to Macalister’s impression (p. 24), the underlying singular suffix does not carry gender distinction. Rather, a vowel assimilation rule renders it as *-ēk* following *-a-* and as *-ik* following *-i-*. Hence masculine nouns and adjectives ending in *-a* take *-ēk*, but also feminines in *-a*, cf. *dōmiya* ‘a Dom woman’, *ihī dōmiyēk* ‘this is a Dom woman’, while feminine nouns and adjectives in *-i* take *-ik*, but also masculines in *-i*, cf. *tmali* ‘soldier’, *pandži tmalik* ‘he is a soldier’. Consonantal stems take a different set, namely *-i/-ī* (sg) and *-ni* (pl): *ihī dīrōm-ī* ‘this is my daughter’, *mušš giš dōme ēkak-ī* ‘not all Dom are alike (= one)’, *pandžan taṣbān-ni* ‘they are tired’. The ending in *-i/-ī* has been interpreted by Macalister as a case of phonological variation, and elsewhere (p. 9) as an ending attached to Arabic loans. This however is clearly contested by the consistency in the distribution of the two sets of predication markers. In presentative constructions we find Indic masculine *šnōṭēk* ‘dog’, alongside Arabic-derived, feminine *zahrēk* ‘flower’ (< *šnōṭa*, *zahra*); and on the other hand *qarī* ‘donkey’, alongside Arabic-derived *qamārī* ‘moon’ (< *qar*, *qamar*). Both these forms appear to have cognates in other Domari varieties; we find in Patkanoff’s (1907/1908) material a *-k* form serving as a 3sg ending on the copula – *asta-q*, and a 3sg existential form in *hi*.

The forms in *ēk/ik/ēni* serve as converbal or present participle endings when attached to the past participle of lexical verbs: *gar-ēk* ‘walking’, *tird-ēk* ‘standing’ (from *tw-/tirda* ‘to place’). With verbs of motion or state, we find frequent use of impersonal predications conveying an accomplished state: *pandži tirdēk* ‘he is standing’. Predication markers form an integral part of the enclitic copula paradigm, taking over third persons: *ama mištahromi* ‘I am ill’, but *pandži mištik* ‘she is ill’, *pandžan mištēni* ‘they are ill’. Occasionally this use of impersonal, non-verbal predication markers is extended to 1st and 2nd persons as well: *ama kuryamēk* ‘I am at

home', *atu kuryamēk* 'you are at home'; *ama dōmī* 'I am Dom', *atu dōmī* 'you are Dom'.

From this, several generalisations can be made about the formation of the existential paradigm (Figures 9–10). Copula forms are enclitic, while independent verbs of existence precede the predication noun: *ama mudīrahromi* 'I am director', but *ama (a)hrom mudīr* 'I became director', *ama homi mudīr* 'I am becoming director'. Present-tense copula forms for the 1st and 2nd persons are based on the underlying perfect form of the existential verb, in enclitic position. For the 3rd person they are based strictly on the impersonal predication markers; this partly also extends to the other persons. The original present forms of the copula are attested only for the 1sg, 2sg, and 3pl, and even those appear only marginally.<sup>14</sup>

Figure 9: Attested forms of the present enclitic copula

1sg	-ahromi	-ēk/-ik, -ī	-homi
2sg	-ahrōri	-ēk/-ik, -ī	-hōk
3sg		-ēk/-ik, -ī	(-hi)
1pl	-ahrēni	-ēk, -nī	
2pl	-ahrēsi	-ēk, -nī	
3pl		-ēnī, -nī	-hodi

Figure 10: The independent existential verb 'to become'

	Present	Subjunctive/Imper.	Perfective/ -Perfect
1sg	<i>hōmi</i>	<i>hōšam</i>	(a)hrom -i
2sg	<i>hōšēk</i>	<i>hōšī</i>	(a)hror -i
3sg m/f	<i>hōsari</i>	<i>hōšar</i>	(a)hra / (a)hrī
1pl	<i>hōšani</i>	<i>hōšan</i>	(a)hrēn -i
2pl	<i>hōšasi</i>	<i>hōšas</i>	(a)hrēs -i
3pl	<i>hōša(n)di</i>	<i>hōsad</i>	(a)hre

The variation in the copula paradigm applies to the present indicative only, and the subjunctive is *hōšam*, *hōšī* etc. throughout, preserving the stem *hō-* and the subjunctive extension. Since the subjunctive/imperative necessarily conveys a transition and not a state, being employed in modal constructions

14 Macalister (p. 35) documents 1st and 3rd persons, singular and plural, and speculates on the 2nd, suggesting *\*hwēki* and *\*hwēsi*.

only, its position is variable, as enclitic copula and independent existential verb merge: *biddiš mištahōšam /biddiš hōšam mišta* ‘I do not want to become ill’. Noteworthy is the extension of the subjunctive marker *-š-* into the indicative paradigm of the independent existential verb, with the exception of the 1sg (see Figure 10).

The past tense of the independent existential verb follows the pattern for lexical verbs, while the past tense of the enclitic copula is restricted to the imperfect (resembling the pluperfect where the present copula is based on the perfect form): *knēn ahrora?* ‘where were you?’. As in lexical verbs, a conjugated form of the Arabic past-tense copula *kān-* may reinforce the construction: *kunt kaṣṭaḥroma* ‘I was small’. Only for the 3rd person do we find the remoteness marker *-a* acting independently as a past-tense to the impersonal predicative marker: *pandži tmalik/tmaliya* ‘he is/was a policeman’, *pandži ktīr /ktīra* ‘he is/was a Christian’, *pandži xudžoti kuryamēya* alongside *kuryamahreya* ‘he was at home yesterday’.

### 5.5 Copula-based formation of Arabic loan verbs

The integration pattern for Arabic loan verbs consists of the attachment of reduced forms of the Arabic verb, which Littmann (1920: 132) identifies as the imperative form,<sup>15</sup> to indigenous carrier verbs, which carry the verb inflection. There are two main carrier verbs, *k-* (from *kar-* ‘to do’, occasionally also in its long form), and *(h)o-* ‘to be’. In this Domari belongs to a group of languages under the influence of Arabic in a geographical continuum comprising northern New Indo-Aryan, Iranian, and Turkic. The assignment of Arabic roots in these languages to one of the two carrier verbs – the ‘do’ class and the ‘be/become’ class – tends to follow the transitive/intransitive distinction. In Domari, the ‘do’ class clearly represents an older layer of carrier verbs. This can be seen first in the overwhelming tendency to use a shortened, grammaticalised form of the verb ‘to do’ as a carrier, and second in the employment of *k-* also with pre-Arabic items, among them Indic elements: *mangišk-* ‘to beg’, *kamk-* ‘to

15 This is not entirely unproblematic, however. Consider *ʕazifk-* ‘to play music’, which has an Arabic imperative in *(i)ʕzaf*. The discrepancy can only be resolved by assuming, as Littmann (1920: 131) does, a movement between Arabic verb derivation classes. This movement is in turn accompanied by morphophonological changes, which are significant, since they actually violate the general pattern of verb class formation. Thus, in the case of *ʕazifk-*, the root *ʕzf* would be imitating either the second class, but dropping gemination of the second consonant, or alternatively copying the third class, while ignoring lengthening of the first vowel.

work', *ladžik*- 'to be shy'. Of the Arabic loans in the 'do'-class, those encountered in the corpus so far are indeed transitives.

Problematic in the context of the present discussion is the class of 'be'-based loan verbs. The class appears to be open to both intransitives and unergatives of the type *fhim*- 'to understand', *hibb*- 'to love', *iḥtiram*- 'to respect'. It is evidently younger than the 'do'-class, consisting entirely of Arabic-derived lexical items. Structurally, it is affected by the still ongoing transition of perfect forms of the copula to a new present paradigm, as well as in part by the infiltration into this paradigm of impersonal converbal endings, for the 3rd person.

Figure 11: Patterns for copula-based loan verbs  
(Subjunctive always with *hōš*-; Imperfect with *kān*- and optionally *-a*)

	Pattern 1 <i>fhim</i> - 'to understand'	Pattern 2 ʕīš- 'to live'	Pattern 3 <i>tšib</i> - 'to be tired'
Present			
1sg	<i>fhimomi</i>	ʕīšahromi	<i>tašbānahromi</i>
2sg	<i>fhimōk</i>	ʕīšahrori	<i>tašbānahrori</i>
3sg	<i>fhimori</i> , <i>fhimrēk/ -rik</i>	ʕīšahrēk/ -rik	<i>tašbānī/ -ik</i>
1pl	<i>fhimoni</i>	ʕīšahrēni	<i>tašbānahrēni</i>
2pl	<i>fhimosi</i>	ʕīšahrēsi	<i>tašbānahrēsi</i>
3pl	<i>fhimodi</i>	ʕīšahrē(n)di	<i>tašbānnī</i>
Perfective			
1sg	<i>fhim(ah)rom</i>	ʕīšrom	<i>tšibrom</i>
2sg	<i>fhim(ah)ror</i>	ʕīšror	<i>tšibror</i>
3sg	<i>fhim(ah)ra/ -ī</i>	ʕīšra/ -ī	<i>tšibra/ -ī</i>
1pl	<i>fhim(ah)rēn</i>	ʕīšrēn	<i>tšibrēn</i>
2pl	<i>fhim(ah)rēs</i>	ʕīšrēs	<i>tšibrēs</i>
3pl	<i>fhim(ah)re</i>	ʕīšre	<i>tšibre</i>

Three structural patterns can be identified for the formation of present and past tenses of the 'be'-class of Arabic loans (Figure 11). They are only partly complementary, as there is some alternation among categories for individual verbs, as well as variation both among speakers and within the speech of individual speakers. The patterns are clearly motivated by an effort to avoid ambiguity which might arise as a result of the movement

across tense formations, with underlying perfect forms being interpreted potentially as either present or past. Pattern 1 draws on the original present paradigm of the copula, normally deleting the *h*- stem, for the present tense. For the past it employs either the perfective form of the existential verb in (*a*)*hr*-, or else simply a perfective extension in *-r*-. Pattern 2 draws on the renewed copula, based on the perfect of the existential verb, for the present tense, and employs for the past the bare perfective extension in *-r*-.

Note that the historical division among consonantal stems and vowel stems, which originally conditioned the distribution of the perfective extensions (underlying participle in *-ta*) in *-d*- and *-r*- respectively, is herewith dissolved, as Arabic-derived stems such as *skunn*- 'to reside' or *tšallim*- 'to learn' take an *-r*- extension, owing to the existential verb figuring as carrier. The overall synchronic division between transitives (*-d*-) and intransitives/unergatives (*-r*-) remains, however. Pattern 3 finally comprises those verbs whose present tense draws on an Arabic present participle, to which the (renewed) Domari enclitic copula is attached, while the past is formed on the basis of the reduced verbal stem, with the perfective extension likewise in *-r*-.

## 6. COMPLEX CLAUSES

The use in Domari of an Arabic frame for combining clauses, including Arabic conjunctions, as well as Arabic rules and patterns of word order has already been noted by Littmann (1920) as well as Macalister. It is worth stressing nevertheless that Domari does not simply borrow individual particles and conjunctions from Arabic, but that, rather, Arabic structures constitute the only overt means for clause combining. Domari may thus be said to have undergone complete 'fusion' (see Matras 1998) with Arabic in its clause combining structures. Exempted from this are only the most tightly-integrated clauses on a hierarchy of clause integration (cf. Givón 1990). The present discussion will survey some positions along this hierarchy.

### 6.1 Embeddings and relative clauses

Embeddings, where the subordinated clause is a constituent, show the only use of indigenous *wh*-elements in subordinated clauses:

- (5) a. *ama džaname ' krēn skunahrēk*  
       I know.1sg.neg where live.pred  
       'I don't know where he lives'



- b. *džandomi kī širda*  
 knew.1sg.perf what said.3sg.m  
 'I understood what he said'

Isolated examples in the corpus illustrate nevertheless the beginning infiltration of Arabic structures even here: *džaname' 'ēš biddī karam* 'I don't know what I want to do' (< Arabic 'ēš).

Relative clauses are introduced through the Arabic relativiser 'illi. Like Arabic, Domari too has an obligatory resumptive pronoun for all positions except the subject. Where only one object appears, resumption of the head noun is indicated through object pronominal clitics on the verb (direct object) or a location expression (indirect object):

- (6) a. *kažža 'illi laherdomis xužoti*  
 man rel saw.1sg.3sg yesterday  
 'the man whom I saw [him] yesterday'  
 b. *lāši 'illi širdom wāsīs*  
 girl rel spoke.1sg with.3sg  
 'the girl whom I spoke to [her]'

In principle the same strategy may be followed when the relative clause contains two objects:

- (7) a. *mana 'illi toris amake*  
 bread rel gave.2sg.3sg me.ben  
 'the bread which you gave [it] to me'  
 b. *ple 'illi torsan amake*  
 money rel gave.2sg.3pl me.bef  
 'the money (pl) which you gave [them] to me'

At the same time there is also a tendency to employ an Arabic resumptive pronoun for a head noun that is the direct object of the relative clause, while the indirect object is expressed as a pronominal clitic on the verb. Arabic inflection is then used to mark agreement in gender and number between the Arabic resumptive pronoun and its Domari head noun:

- (8) a. *mana 'illi torim iyyāh*  
 bread rel gave.2sg.1sg res.3sg  
 'the bread which you gave [me] it'  
 b. *ple 'illi torim iyyāhum*  
 money rel gave.2sg.1sg res.3pl  
 'the money (pl) which you gave [me] them'

Note that this has a double effect on the expression of syntactic relations within the sentence: Firstly, what is generally marked as an indirect object, namely the benefactive of the verb ‘to give’, is expressed as a pronominal clitic using the set of markers and the position in the verb normally reserved for direct objects. Second, Arabic inflection is used productively within the Domari sentence. I shall return to this latter point briefly below.

## 6.2 Complementation and purpose clauses

The key features of complementation structures in Domari are the split between subjunctive and indicative complements (indicated in Domari, as in Arabic, through the choice of mood in the subordinated clause), the presence of a conjunction, and the choice of a modal expression that requires a modal complement. There are only two modal expressions in Domari that are inherited: *sak-* ‘to be able to’, and *mang-* ‘to ask’, which latter is restricted to different-subject modal constructions (manipulation). Other modal expressions are Arabic, and carry, if inflected, Arabic inflections: *lāzim* ‘must’ (impersonal), *ṣār-* ‘to be begin’ (inflected), *bidd-* ‘to want’ (nominal inflection), *xallī-* ‘to allow’ (inflected).

As in Arabic, with same-subject modality no complementiser appears between the main and the complement clause, and the subordinated verb is finite and subjunctive (*biddī karam* ‘I want to do’). Manipulation clauses equally require no conjunction, but an overt representation of the manipulee must be present; the subordinated verb is likewise in the subjunctive:

- (9) a. *ama mangedom minšīs šrīkar mana*  
 I asked.1sg from.3sg buy.3sg.subj bread  
 ‘I asked him to buy bread’  
 b. *ama širdom abuske awar wāšīm*  
 I said.1sg to.3sg.ben come.3sg.subj with.1sg  
 ‘I told him to come with me’  
 c. *ama biddī atu šrīka manas*  
 I want.1sg you buy.2sg.subj bread.acc  
 ‘I want you to buy the bread’

In purpose clauses there is variation in the presence vs absence of a conjunction (which is, if present, always Arabic-derived). The split may be said to follow a continuum of semantic integration, or in some instances, control by the main actor over the action conveyed by the purpose clause, thus resembling the distribution in Arabic:

- (10) a. *nan frayēm warkamsan!*  
bring clothes.pl.1sg wear.1sg.subj.3pl  
'bring my clothes for me to wear'
- b. *ama ērom kuryata (šašān) lahamir*  
I came.1sg house.dat comp see.1sg.subj.2sg  
'I came home to see you'
- c. *ama tomir ple (šašān) šrika mana*  
I gave.1sg.2sg money comp buy.2sg.subj bread  
'I gave you money to buy bread'
- d. *ama qoldom qapiya šašān [\*ø] nikšī*  
I opened.1sg door.acc comp enter.2sg.subj  
'I opened the door so that you may enter'

Indicative complements follow epistemic verbs. The subordinated verb is in the indicative, and the complement is always introduced by an (Arabic) conjunction *inn-*, which may assume either an impersonal or an inflected form (carrying Arabic inflection):

- (11) *ama sindom inn-o/inn-ak atu šīšrori hinēn*  
I heard.1sg comp-3sg/comp-2sg you lived.2sg here  
'I heard that you live(d) here'

### 6.3 Adverbial clauses

Domari has converbs which express a co-occurring action. They are based on the attachment of the predicative suffix to the perfective form of the verb (12a–b). The same function however can also be assumed by the finite present form of the verb (12c), once again matching Arabic, which has two options, present participle and present/future, to express simultaneous action:

- (12) a. *lakedomis mindirdēk*  
saw.1sg.3sg stand.perf.pred  
'I saw him standing'
- b. *lakedom qapiya banīrik*  
saw.1sg door.acc open.perf.pred  
'I saw the door opening'
- c. *sindomis grēfkari*  
heard.1sg.3sg sing.3sg  
'I heard him singing'

Other adverbial subordinations draw on Arabic conjunctions. Noteworthy is the distribution of tense and mood forms: Anteriority (13a) triggers, as in Arabic, the subjunctive in the subordinated clause. Realis conditionals show the present tense in both parts of the construction (13b). Irrealis (counterfactual) constructions have unspecified perfective in the subordinated clause, and pluperfect, introduced by the Arabic particle *kān*, in the main clause (13c; see also above):

- (13) a. *qabel mā džam xatṭaškedom kamas*  
 before comp go.1sg.subj finished.1sg work.acc  
 ‘Before I left I finished my work’  
 b. *iza warsari, nawame’*  
 if rain.3sg neg.come.1sg.neg  
 ‘If it rains, I shall not come’  
 c. *law ėrom xužoti kān laherdomsa*  
 if came.1sg yesterday was saw.1sg.3sg.pluperf  
 ‘If I had come yesterday, I would have seen him’

## 7. THE ARABIC COMPONENT

From section 6 it is clear that Arabic has had a considerable impact on sentence structure and especially the grammar of clause linkage in Domari. Indeed, it is only in the most tightly-integrated clauses (embeddings, adverbial simultaneity, modality) that no overt Arabic clause-linking devices are employed, though even here the two languages may be said to be compatible, and the absence of conjunctions can be taken to reflect the universal rules on syntactic integration of semantically closely-linked clauses. Littmann (1920) has already furnished a rather thorough description of the Arabic component of Domari, based on Macalister’s published material. It includes a discussion of phonological and lexical features, patterns of verb integration (see also discussion above), word order rules, and more. I choose only to add two basic remarks in the present framework; both pertain to the categorial status of the Arabic component in the Domari grammatical system.

The first remark concerns the distinction between borrowing of structures, convergence, and fusion. Borrowing, defined as the adoption of Arabic-derived items, occurs in virtually all grammatical domains, but is particularly conspicuous in the lexicon. Convergence, defined as the establishment of structural compatibility among the languages even in the

absence of actual shared material, typically concerns morphosyntactic areas where the distribution of Domari structures matches that of their Arabic counterparts. This can be said to apply to the order of constituents in the sentence. Arabic word order is still largely resisted within the noun phrase, with adjectives generally preceding the noun, though single cases of noun-adjective order, noted already by Littmann (1920: 136) based on Macalister's texts, appear in the corpus. Convergence also applies to the distribution of moods and the distinction between indicative and subjunctive complements. Partly we find convergence in the distribution of tenses, notably in the occurrence of non-finite forms (predicative suffix forms, matching Arabic present participles), and the use of tenses in adverbial clauses. Domari of course possesses tenses that are not present in Arabic, namely perfect and pluperfect, and it is necessary to correct the impression given by Littmann of a complete match in tense distribution (based on Macalister's description of tenses).

We come to fusion. I have defined fusion (Matras 1998) as the non-separation of languages for a particular grammatical category, resulting in the wholesale adoption of a class of items from language B, replacing its indigenous counterparts in language A. Fusion categories in Domari include all interaction-managing elements such as discourse particles, focus particles, sentential and phasal adverbs, and overt clause combining devices. Domari thus follows the prediction that fusion will begin with utterance-modifying elements, also expressed in Stolz & Stolz's (1996) hierarchy of grammatical borrowing which predicts earlier transfer of discourse-level operators. To these universally predictable cases of fusion we must add the class of prepositions. Inherited adpositions remain in Domari only as person-inflected local relation expressions, which in essence function as case-marked personal pronouns. All genuine adpositions are Arabic-derived. The emergence of Arabic prepositions can be explained as the outcome of a general typological shift as a result of which adpositions assume a preposed position, while at the same time the productivity of inherited material weakens to such extent, that the language is unable to form its own prepositions from an inherited stock of location expressions.

The second remark on the Arabic component concerns the productivity of Arabic inflection. This is encountered in a series of classes. Both Macalister and Littmann comment on the occasional presence of an Arabic

definite article accompanying Arabic lexical material; consider the following example from the corpus:

- (14) *misilmīne nḥibbodmané', řaşān ihne' ama nmangamsané'*  
 Muslims.pl neg.like.3pl.1pl.neg because such I neg.like.1sg.3pl.neg  
*l-řarab*  
 def-Arabs  
 'The Muslims don't like us, that is why I don't like the Arabs'

Note that the word *misilm-īn-e* also carries a double plural marking, adding the Domari ending to the Arabic one. The presence of the Arabic article and nominal derivation reminds us of the incorporation of noun phrase features from the donor language of lexical material in a mixed language like Michif (Bakker 1997), where French nouns in a Cree clause structure retain French grammatical markers. In Domari, one cannot speak of a systematic or consistent incorporation of the Arabic article. Nonetheless the admissibility of an Arabic article has further implications for syntactic constructions in the language. Thus we find genitive-possessive constructions where the Arabic head assumes a form identical to its form in an Arabic genitive-possessive construction, resulting in a synchronisation of the counterpart structures of the two languages:

- (15) *fī dēyak min dēyeskī l-řīrāq*  
 in town.indef from town.pl.poss.abl def-iraq  
 'In one of the Iraqi towns'

Note the overall extent of hybrid structure here: The idiom itself – *an X from among the Xs* meaning 'a certain X' – is borrowed from Arabic. The prepositions which carry the idiom, defining the categorisation of the participating nouns, are Arabic-derived. Finally the possessor-determiner assumes an Arabic form, and does not show Domari ablative marking which is normally required. Rather, its syntactic role is marked solely, as in Arabic, by its position in the noun phrase.

Arabic inflection proper is encountered (i) in the subject and object inflection of modals, carrying Arabic object/possessive pronominal clitics (*biddī, biddak* 'I want, you want' etc.; *xallīh, xallīha* 'let him, let her' etc.), (ii) in the person inflection and negation of the auxiliary verbs *řār* 'to begin' and *kān* 'to be' (Macalister, p. 35, notes that the form is uninflected), (iii) in the person inflection expressing head agreement in the resumptive pronoun, and different subject agreement in complementisers.

The implications of such extensive presence of productive Arabic inflection are twofold. Firstly, the two systems, Domari and Arabic, are synchronised to allow cross-reference and agreement between them throughout the sentence (Arabic items are glossed in square brackets):

- (16) a. *ama mā kuntiš kuryamēk*  
 I [neg was.1sg.neg] home.loc.pred  
 'I was not at home'  
 b. *xallīhum skunnhōšad barariyama*  
 [let.3pl] live.subj.3pl outside.loc  
 'Let them live outdoors'  
 c. *ama sindom innak atu mnēn hrori*  
 I heard.1sg [that.2sg] you here are.2sg  
 'I heard that you are here'

The integration of Arabic inflection also introduces grammatical distinctions that are not present in the Domari system, notably certain types of gender agreement:

- (17) a. *bašdēn kānat šara amake biddhā qumnar*  
 [then] [was.3sg.f] say.3sg.imp me.ben [wants.3sg.f] eat.3sg.subj  
 'Then she would say to me that she wants to eat'  
 b. *na kildom bara li' annhā warsari*  
 neg went.out.1sg out [because.3sg.f] rain.3sg  
 'I did not go out because it was raining'

In (17a), the feminine gender of the subject is not indicated on the Domari verb, nor would a personal pronoun include such indication, but it is marked on the Arabic verb and modal. In (17b), the complementiser agrees with what in Arabic would be the feminine subject of the verb 'to rain', namely *ad-dunyā* 'the world/nature', which is in turn interpreted into the Domari verb. The two examples suggest that for a synthesis of inflectional patterns to function, sentence processing must take place simultaneously in Domari and Arabic, applying complementary parameters of cross-reference operations to the sentential constituents.

## 8. LINGUISTIC STRATIFICATION

A thorough sociolinguistic survey of the Domari-speaking community is beyond the scope of the present work, but I will provide a brief introductory sketch. It is evident that in most families Domari gradually ceased to

be the principal language spoken to children during the early 1970s. This is the period following the Israeli occupation, the emigration of a sizeable part of the community to Jordan, the introduction of free compulsory elementary education (in Arabic), and the completed transition to paid employment outside the community. Though a connection must be assumed between these factors of social change in the Jerusalem community and the decline of the language, there is reason to believe that Domari is in a comparable state elsewhere as well. I base this impression on conversations with Doms from Amman and Gaza, and on the remarks made in Meyer (1994) with regard to the Dom population in and around Damascus.

Speakers with any degree of familiarity with the language can be divided into roughly four strata. Two of those comprise actual fluent speakers of the language, the other two include what one might call semi-speakers. Among the fluent speakers, Group A consists of an older generation of speakers, more specifically however of consistent users of a cluster of more conservative linguistic features. Group B, on the other hand, are usually somewhat younger, and show a consistent clustering of more recent innovations.

Figure 12: Selected differences among two strata of fluent speakers

	Group A ('older')	Group B ('younger')
instrumental marker	<i>-san</i>	<i>maʕ ... kī</i>
ablative marker	<i>-kī</i>	<i>min ... kī</i>
dative marker	<i>-ta</i>	<i>-ka</i>
3pl pres verb concord	<i>-ndi</i>	<i>-di</i>
purpose clause COMP <Ar.	<i>ta</i>	<i>ʕaʕān</i>
phonetics	č-; -ū- in possessives	š-; -o- in possessives
NP retrieval ('small boy')	isolated/nom.: <i>kaštota zara</i>	predicative: <i>zarēk kaštotēk</i>

The most salient features are summarised in Figure 12. They include modifications to the case system, either through reliance on Arabic material or, in the case of the dative marker, internal change; simplification of the 3pl concord marker; use of a less conservative Arabic purpose conjunction (coinciding with a change in Jerusalem Arabic), phonetic innovations; and finally the strategy for structural retrieval and presentation of isolated noun phrases, noticeable especially during questionnaire elicitation. Here, conser-



vative speakers will cite nominative forms, while Group B speakers will form non-verbal predicative constructions.

Some of these changes were already pointed out by Macalister when discussing variation in the language (see remarks in the preceding sections). Interestingly, the speakers interviewed for the present study tend to favour particular variants consistently. Moreover, the clustering of variants as described in Figure 12 is also consistent, at least in the data collected so far. Group A speakers have in addition a considerably larger Domari vocabulary at their active disposal, while Group B speakers will use more Arabisms. (One should note a general decline in speakers' command of vocabulary, compared to the stage documented by Macalister; this includes numerals, knowledge and use of which is now restricted to those up to 'five', while Macalister was able to document the complete set). All this allows us to identify two distinct codes among the fluent speakers, some of whom are very closely related through family ties, and all of whom live in close proximity to each other and interact almost daily. Beyond tentative age, no clear extra-linguistic correlates for the use of these feature clusters or 'codes' could so far be identified.

The strata of semi-speakers can likewise be divided into two Groups, C and D. Group C speakers, in their twenties and thirties, can be considered to possess a good passive knowledge of Domari, and are able on occasion to produce sentences, at least when asked to do so, although they would normally not use Domari actively. In such instances, Group C speakers often confuse or neglect altogether conjugation endings (18a,b). Noun inflection is similarly eroded (18b,c,d): Case is replaced entirely by Arabic prepositions, word order in the noun phrase shows Arabic noun-adjective (18c), and mixing with Arabic extends beyond the 'permissible' degree – e.g. we find in (18d) Arabic articles with indigenous nouns, and Arabic tense inflections with lexical verbs:

- (18) a. *itme biddkū bīsawahōšad*  
           you.pl [want.2pl] marry.subj.3pl  
           'You want to marry' (verb fails to agree with subject)
- b. *žanami aha kažža*  
           know.1sg this man.nom  
           'I know this man' (direct object in the nominative instead of accusative)

- c. *skunnahrom bi kurya tilli*  
 live.1sg [in] house.acc big.f  
 'I live in a big house' (noun-adj word order)
- d. *šar abuske šan il-lāšyēnī bitžawwazka mišš*  
 tell 3sg.ben [about] [def]girls.pred [pres.3pl]marry.ø [not]  
*dōme*  
 dom.nom.pl  
 'Tell him about the girls who marry non-Doms'

Groups D comprises those, usually in their late teens or early twenties, who are only able to retrieve a knowledge of single, isolated lexical items. Noteworthy is the fact that these are never retrieved in isolation. A closed class of expressions of intimate possession (family, parts of the body, home) are always inflected for the 1sg possessive (thus *kuryom* for 'house' < 'my house'). Other items always appear with predicative markers.<sup>16</sup> The impression is one of context-based selective replication of items, with no ability for context-independent disambiguation.

## 9. DOMARI AND ROMANI

The relationship between Domari and Romani has been the focus of linguistic debates since Pott's (1846) discussion of the language of the Syrian Gypsies. The question of linguistic affinity has since been regarded as a key issue in attempts to reconstruct Romani origins, and more recently to shape the historical narrative describing the emergence of Romani identity (see Hancock 1998). The position of ethnographic evidence in the comparison of the two groups is controversial. There is no doubt that the affinity felt by the Doms themselves and their recent interest in the Roms of Europe and America arises not solely on the basis of linguistic similarities, but also through similar status in society, similarities in the collective views concerning their relations with mainstream society, and similarities in the traditional occupation patterns which have been an essential component of both Dom and Rom identity in the past. While none of these may be overlooked when attempting to provide an answer to the question of origins, the present discussion is concerned primarily with linguistic aspects. In the above sections I have already made repeated reference to Romani, and my

16 This strategy is already noted by Macalister (p. 10).

concluding remarks will be devoted to a concise comparison of the two languages.

When comparing Romani and Domari it is of course crucial to differentiate between material that is part of the historical legacy of Indo-Aryan on which both languages ultimately draw, and innovations. Furthermore, a distinction must be sought between different types of innovations. There are those innovations that follow a natural track of development, those which we might, in other words, expect a language to follow given our knowledge about a particular point of departure at a given stage in its history. Such is the attachment of Layer II case markers to the noun, exhibited to various degrees by NIA languages. Once a small and closed set of postposed, abstract location markers have become grammaticalised, they can either follow closely after the noun, or show various degrees of integration. Another kind of innovation may be triggered by language contact. Here, similarities between the two languages do not necessarily reflect a close historical affinity. Consider for example the loss of the infinitive in modal constructions. In Romani this is typically considered a Balkanism, while in Domari it can be explained as an outcome of Persian, Kurdish, or Arabic influence. Naturally, differences between the languages that can be explained by contact do not a priori contradict the possibility of a shared origin, either. Consider the emergence of a definite article based on the Greek model in Romani, which is lacking in Domari. This is clearly a late envelopment that will have occurred long after a split of the two branches, should we assume that they descend from a common linguistic unit. Our primary task is therefore to keep apart (i) the various paths of natural selection from a shared inherited stock, (ii) innovations that are a result of diverse contact constellations following emigration from India, and (iii) innovations that are not entirely predictable, and especially clusters of non-predictable, yet shared innovations that might point to a contiguous development.

Let us first examine natural selection from the shared inherited stock – the OIA and MIA legacy. On the whole Domari and Romani share the bulk of their basic vocabulary, as demonstrated by a comparison of the entries for the hundred word list (see Hancock 1988, based on Sampson 1926). Domari items that differ from Romani can, however, also fall into the domain of primary vocabulary; since we must assume early multilingualism in ever-changing constellations for both communities, given the general volatility of vocabulary, the non-compatibility of portions of the

basic vocabulary should not in itself exclude a close affinity of the two languages in an earlier period. The fact that regular sound correspondences can explain the forms of most cognates in the two languages is merely an attestation of vocabulary drawing on a general Indo-Aryan inherited stock. There is little evidence of shared phonological innovation that distinguishes the two from other NIA languages, save perhaps the development of dentals into liquids (Domari *r*, Romani *l*: MIA *dža-ti* > D *dža-ri*, R *dža-la*) (cf. Turner 1926). The loss of retroflexes is a predictable contact-related change in each of the languages.

In morphology, both Domari and Romani are conservative in preserving much of the MIA present verb conjugation as well as the consonantal forms for Layer I markers. The forms employed for Layer II affixes are shared for some functions, namely the Benefactive (*ke*), Dative (*ta*), and Sociative (*san*). In addition, a connection may be drawn between the Romani Genitive and the Domari Ablative (*k-/kī*). No cognates exist in Domari for the Romani Ablative (*tar*) and in Romani for the Domari Locative (*ma*). Interrogative forms are likewise only partly shared. Domari relies entirely on the *k*-forms, while Romani shows a set of forms in *s*- (*so*, *sar*, *savo* ‘what, how, which’). Location adverbs display a similar picture: shared are D. *ager/ab/pač/bar*, R. *a(n)gl(a)p/paš/(a)vr* ‘in front/on/behind/out’, distinct are D. *atun/axar/mandž*, R. *opr-/tel-/and-* ‘above/below/inside’. Domari deictic forms share the vowel opposition *e/o* with their Romani counterparts (*a/o*), while the postvocalic stem reflecting etymological *-ta* in oblique deictics is realised as *-r* in Domari (*oras*) and as *-l* in Romani (*oles*); otherwise Domari does not share the multiplicity of Romani deictic stems in *d-/v-/j-/k-*. The *h*- deictics present in Domari have been compared to the vocalic definite article in Romani (*aha*, *ihi*, *ehe* > *o*, *i*, *e*; see Sampson 1926: 152).

Auxiliary verbs present a complex picture. For the existential verb Domari has forms in *št-* and in *h-* (the *št-* forms are more productive in northern varieties, and appear in Jerusalem Domari only in the impersonal *ašte* ‘there is’). Romani has forms in *s-* and in *h-*. It is not clear whether these reflect different lexical verbs – OIA *as-*, *bhū-*, and perhaps also *sih-* – or phonological alternations within the *as*-derived paradigm, where MIA had optional aspiration of *s* > *h*. A distinct Domari development is the sub-

junctive suffix *-š-* < *-č-*, likely to be derived from *aččh-* ‘to stay’.<sup>17</sup> This is found as a copula auxiliary especially in eastern, but also in other NIA languages. In Domari it appears to have been confined initially to verbs expressing state and motion. In this connection we can also note the retention in Domari of the OIA/MIA passive derivation morpheme in *-y-*; Romani, on the other hand, grammaticises a passive auxiliary in *\*ov-* which eventually becomes integrated into the synthetic morphology of the verb.

In sum, then, we must conclude that natural selection among inherited forms does not point to an identical or even to a strikingly close development of the two languages. We now turn briefly to contact-related innovations. As pointed out by Hancock (1995), Domari and Romani do not share most of their vocabulary of Iranian origin, and so even if one should assume a shared origin within India, separation in or before migration through Iranian-speaking territory, which is what Sampson (1923) had argued for, is likely. Perhaps the most significant contact-related innovations that are shared are the loss of a modal infinitive, the shift to verb-object word order (though not in northern Syrian Domari, where verb-final order prevails), and the presence of prepositions. For Romani, it is attractive to regard these developments as part of the Balkanisation process, while for Domari they may have been triggered in part through contact with Iranian, and are synchronically definitely reinforced through convergence with Arabic. Distinct innovations that are likely to be contact-related include the aforementioned preposed definite article in Romani, the postposed indefinite article in Domari (possibly a result of Kurdish influence), and the enclitic position of the Domari copula, reinforced perhaps by the enclitic position of the copula in Iranian and Turkic, and further by the lack altogether of a present-tense copula in Arabic, which leaves Domari no model to copy. Distinct typologically are the use in Domari of predicative suffixes for non-verbal predications, the presence in Domari of a converbal form, and the distinct genitive-possessive constructions employed by the two languages. Finally we note the productivity of items of inherited stock for the formation of prepositions and conjunctions in Romani, but not in Domari, which undergoes fusion with Arabic in these domains.

17 Barr (1943: 45–46) however regards this as Iranian influence, and specifically as a combination of the Persian prefix *ha-* with the Kurdish verb stem *č* ‘to go’. I do not find this etymology very convincing.

We are left with a series of internal innovations. The most conspicuous features shared by Domari and Romani are the synthetisation of Layer II case markers, and the emergence of subject concord markers on the past form of the verb. The first development might be regarded as a natural continuation of the process of attachment of abstract markers to the noun phrase, mediated by Layer I. Such markers may either remain detachable from the noun, as in Hindi, or become integrated, as in Bengali, Domari, and Romani. If we regard synthetisation of case affixes as a continuum, however, we discover that the developments in Domari and Romani are not identical. In Domari, Layer II affixes may be separated from the noun through personal possessive markers. Moreover, unlike Romani, there is no phonological assimilation to preceding consonants and so a weaker degree of structural integration.

The second development, the emergence of a past conjugation, can be seen in the general context of the attachment of a 'be' auxiliary to the participle form. This is encountered in other NIA languages as well, though it does not always lead to synthetisation and to the emergence of a new conjugation. Once more we have no direct evidence that both languages underwent the development together; clearly, both Romani and Domari can be said to lack the resistance toward synthetisation – in both case markers and verb conjugation – that is characteristic of the northwest-central NIA languages. The choice of similar person affixes – *-om* in the 1sg.; D. *-r*, R. *-l* in the 2sg; and adjectival endings (intransitive) alternating with *-s* (transitive) in the 3sg. – might be viewed as evidence in support of a shared development. The regularity of the phonological development of post-vocalic *t* to a liquid in both languages (Domari *r*, Romani *l*), as opposed to the voicing of postdental *t*, renders a similar split in the past tense formation: D. *garom*, *kardom* R. *gelom*, *kedom* 'I went, I did'. A final shared innovation involves the attachment of a remoteness marker – Domari *-a*, Romani *-as* – to the verb in external (right-most) position, forming the imperfect from the present, and the pluperfect from the unspecified perfective.

From this one may postulate a shared cluster of innovation-based isoglosses comprising the following features: (i) copula auxiliaries make use of cognate concord markers, at least in the singular, (ii) these auxiliaries undergo synthetic attachment to the participle stem, giving rise to a new set of subject concord markers, (iii) a similar phonological development of *t* results in a similar split within the past-tense paradigm, (iv) there is external tense formation in both languages through affixation of a remoteness

marker, and (v) the two languages appear in neighbouring, though not in identical positions on the continuum toward synthetisation of Layer II case markers.

What about internal innovations that separate the two idioms? As mentioned, the roots of the variation in the copula paradigm are not sufficiently clear at this stage; a conventionalisation of accented and de-accented copula forms in *s* and *h* respectively, as suggested by Boretzky (1995), cannot be ruled out, though nor can the grammaticalisation of distinct lexical verbs. It appears that Domari and Romani at some stage had similar if not identical auxiliary copula forms, which served as concord and tense carriers for past-tense lexical verbs. A distinct development can be reconstructed nonetheless, with the Romani independent present copula continuing to show the same set of concord markers as the auxiliary (now the past conjugation), while the past tense of the copula is formed through addition of the remoteness marker *-as* (*som-somas* 'I am-was'). In Domari, on the other hand, the copula follows the structure of lexical verbs. Its present conjugation matches that of lexical verbs, and its past tense is formed by attachment of the perfective extension in *-r-* (*ho-m* > \* *h(o)r-om* > *ahr-om* 'I am > was'; *hō-k* > \* *h(o)r-or* > *ahr-or* 'you are > were', etc.). The Romani passive formation mentioned above is a further case for a language-particular innovation.

A rather radical difference is the emergence in Domari of pronominal object clitics. On the noun, they correspond functionally to Romani independent possessive pronouns (D. *dāy-om*, R. *miri daj* 'my mother'). Attached to the verb, they might be taken to document a stage following the transition to verb-object word order: D. *sindom-is*, R. *šundom les* 'I heard him', D. *sindosman*, R. *šundas amen* 's/he heard us'. This may indicate a rather late development. Moreover, the fact that gender distinction in the 3sg is lost in the set of pronominal object clitics, although it is maintained elsewhere in the system, might suggest that cliticisation emerged as a result of convergence with the Persian-southern Kurdish continuum of Iranian dialects – which lack pronominal gender distinction – at a rather late stage.<sup>18</sup> This would allow us to place Domari pronominal clitics among the

- 18 A possible counter-argument is the placement of pronominal clitics in internal position relative to (i) external tense markers *-i* and *-a* (cf. *laham-r-i* 'I see you', *lahedom-r-a* 'I had seen you') and (ii) Layer II case affixes (*bay-im-ke* 'for my father'). The first instance appears however to be a genuine case of tense affixes yielding their position to pronominal clitics, for it concerns even the present tense

contact-induced typological innovations, thereby reducing considerably the relevance of internal developments that lead to distinct innovations.

In conclusion, the survey taken here supports the view that there is little evidence for a shared development of Domari and Romani outside of India, save in the continuation of a shared course of development that had begun earlier. In many cases, contact influences lead to different outcomes. In particular, Jerusalem Domari shows a considerable weakening of its inherited resources and draws heavily on Arabic for syntactic restructuring. The spread of non-verbal predications, the enclitic copula, pronominal object clitics, and the head-possessor construction can be judged to be contact-induced or contact-inspired innovations that result in typological formations that are very distinct from those of Romani.

It is however possible that Domari and Romani shared a period of contiguous existence in India, or perhaps even in a non-Indic speaking environment during a very early period. This would help explain the cluster of internal innovations that they share; however, such an explanation is not absolutely necessary if we take the view that the most important isoglosses which they share – synthetisation of Layer II markers, of copula auxiliaries with the past participle, and of external tense markers; and a phonological split of historical *t* into dental and liquid – all represent possible, predictable developments whose clustering may be coincidental, or may indeed be conditioned by factors such as geographical remoteness or social isolation from the bulk of central and northwestern NIA languages, in which these developments are more constrained. Both Domari and Romani could then be considered fringe languages, be it for geographical or social reasons, that share both conservatism and innovative developments with other fringe languages in the extreme north, east, and south of the Indian sub-continent. Considering the differences in natural selection among the inherited stock of morphological and partly lexical forms, it seems likely that even if Domari and Romani did undergo a period of contiguous development at some stage, this development will have brought together what were, to begin with, two distinct albeit related Indo-Aryan idioms.

ending which, assumingly goes back to OIA *-ami* etc. In the case of the second, synthetisation can, as argued above, be assumed to have reached an advanced though not yet complete stage in Domari, when compared with the integration of Layer II affixes in Romani.



One must keep in mind that arguments against a single linguistic origin of Domari and Romani as a distinct branch within NIA need not exclude the possibility that the Doms and the Roms share a socio-ethnic origin within the caste structure of Indian society, going back to the Indian *dom* caste of service-providers, as already suggested over a century ago (Grierson 1888). Such an origin would account for the ethnographic similarities, including shared ethnonyms, and might help explain the motivation of each individual group to emigrate yet still retain inherited forms of socio-economic organisation, specialisations, and traditions, including language. Moreover, it would allow for an accommodation of the linguistic differences, since caste origin need not overlap with geographical origin, while at the same time allowing for periods of contiguous linguistic development prior to the emigration from India, which might have arisen as a result of internal migrations and re-association with populations of similar status and socio-cultural profile.

#### ABBREVIATIONS

abl	ablative
acc	accusative
ben	benefactive
comp	complementiser (Arabic)
dat	dative
def	definite article (Arabic)
f	feminine
imp	imperfect
indef	indefinite article
loc	locative
m	masculine
neg	negator
nom	nominative
perf	perfect
pl	plural
poss	possessive
pred	predication suffix
rel	relativiser (Arabic)
res	resumptive pronoun (Arabic)
sg	singular
subj	subjunctive

## REFERENCES

- AMANOLAH, SEKANDAR & NORBECK, EDWARD 1975. The Luti, an outcaste group of Iran. *Rice University Studies* 61/2: 1–12.
- BAKKER, PETER 1997. 'A language of our own'. *The genesis of Michif – the mixed Cree-French language of the Canadian Métis*. New York.
- BARR, K. 1943. Bidrag til Sigøjnerdialekternes grammatik. II: Iranske elementer i Nūrī. In: *In Memoriam Kr. Sandfeld. Udgivet paa 70-aarsdagen for hans fødsel*. København: 31–46.
- BENNINGHAUS, RÜDIGER 1991. Les Tsiganes de la Turquie orientale. *Etudes Tsiganes* 1991/3: 47–60.
- BORETZKY, NORBERT 1995. Die Entwicklung der Kopula im Romani. *Grazer Linguistische Studien* 43: 1–50.
- BUBENÍK, VÍT 1997. Object doubling in Romani and the Balkan languages. In: Yaron Matras et al. (eds.), *The typology and dialectology of Romani*. Amsterdam: 95–106.
- CANOVA, GIOVANNI 1981. Notte sulle tradizioni zingare in Egitto attraverso la testimonianza di un capo Nawar. *Lacio Drom* 17/6: 4–25.
- DIESSEL, HOLGER 1999. The morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology* 3: 1–49.
- FINCK, FRANZ N. 1907. Die Grundzüge des armenisch-zigeunerischen Sprachbaus. *Journal of the Gypsy Lore Society*, N.S. 1: 34–60.
- FRIEDMAN, VICTOR A. 1991. Case in Romani: new grammar in old affixes. *Journal of the Gypsy Lore Society*, 5th series 1: 85–102.
- GIVÓN, TALMY 1990. *Syntax. A functional-typological introduction*. II. Amsterdam.
- GOBINEAU, A. 1857. Persische Studien. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 11: 689–699.
- GRIERSON, GEORGE A. 1888. Doms, Jats, and the origin of the Gypsies. *Journal of the Gypsy Lore Society* 1: 71–76.
- GROOME, FRANCIS H. 1891. Persian and Syrian Gypsies. *Journal of the Gypsy Lore Society* 2: 21–27.
- HANCOCK, IAN 1988. The development of Romani linguistics. In: Mohammad Ali Jazayeri & Werner Winter (eds.), *Language and cultures. Studies in honor of Edgar C. Polomé*. Berlin: 183–223.
- . 1995. On the migration and affiliation of the Dōmba: Iranian words in Rom, Lom, and Dom Gypsy. In: Yaron Matras (ed.), *Romani in contact. The history, structure and sociology of a language*. Amsterdam: 25–51.
- . 1998. *The Indian origin and westward migration of the Romani people*. Ms. University of Austin, Texas.
- KENRICK, DONALD 1975–1979. Romanies in the Middle East. 1-3. *Roma* 1/3: 5–9; 2/1: 30–36; 3/1: 23–39.
- KRUSE, FR. 1854. *Ulrich Jasper Seetzen's Reisen durch Syrien, Palästina, Phönicien* [...]. II. Berlin.
- LITTMANN, ENNO 1920. *Zigeuner-Arabisch*. Bonn-Leipzig.
- LORIMER, D. L. R. 1939. *The Dūmāki Language*. Nijmegen.

- MACALISTER, R. A. S. 1914. *The language of the Nawar of Zutt, the nomad smiths of Palestine*. (Gypsy Lore Society Monographs 3.) London.
- MASICA, COLIN P. 1990. *The Indo-Aryan Languages*. Cambridge.
- MATRAS, YARON 1995. Verb evidentials and their discourse function in Vlach Romani narratives. In: Yaron Matras (ed.), *Romani in contact. The history, structure, and sociology of a language*. Amsterdam: 95–123.
- . 1997. The typology of case relations and case layer distribution in Romani. In: Yaron Matras et al. (eds.), *The typology and dialectology of Romani*. Amsterdam: 61–93.
- . 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36: 281–331.
- MEYER, FRANK 1994. *Döm und Turkmân in Stadt und Land Damaskus*. Erlangen.
- NEWBOLD, F. R. S. 1856. The Gypsies of Egypt. *Journal of the Royal Asiatic Society of Great Britain and Ireland* 1856: 285–312.
- NICHOLSON, JANE [Undated.] *A fragment of modern Domari*. Unpublished Ms., University of Austin, Texas.
- ORANSKIJ, I. M. 1977. *Fol' klor i jazyk gissarskix Par' ja (Sredneja Azija)*. Moskva.
- PASPATI, ALEXANDRE G. 1870 [1973]. *Études sur les Tchinghianés ou Bohémiens de l' Empire Ottoman*. Osnabrück.
- PATKANOFF, K. P. 1907–1908. Some words on the dialects of the Transcaucasian Gypsies. *Journal of the Gypsy Lore Society*, N.S. 1: 229–257; 2: 246–266, 325–334.
- PAYNE, JOHN R. 1997. The Central Asian Parya. In: Shirin Akiner & Nicholas Sims-Williams (eds.), *Languages and scripts of Central Asia*. London: 144–153.
- PICKETT, DAVID W. & AGOGINO, GEORGE A. 1960. Two legends of the nails of the cross. *Journal of the Gypsy Lore Society*, 3rd series 39: 73–77.
- POTT, AUGUST F. 1844–1845. *Die Zigeuner in Europa und Asien*. Halle.
- . 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*, 5th series 5: 69–95.
- SAMPSON, JOHN 1923. On the origin and early migration of the Gypsies. *Journal of the Gypsy Lore Society*, 3rd series 2: 156–169.
- . 1926 [1968]. *The dialect of the Gypsies of Wales*. Oxford.
- STOLZ, CHRISTEL & STOLZ, THOMAS 1996. Funktionsentlehnung in Mesoamerika. Spanisch-amerindischer Sprachkontakt (Hispanoindiana II). *Sprachtypologie und Universalienforschung* 49: 86–123.
- TURNER, RALPH L. 1926. The position of Romani in Indo-aryan. *Journal of the Gypsy Lore Society*, 3rd series 5: 145–189.
- YANIV, Y. 1980(?). *Ha-Tso'anim bi-Yehuda uvi-Yerushalayim*. [The Gypsies in Judea and in Jerusalem.] Jerusalem.