Aspects of the Early History of Romani

Claus Peter Zoller
University of Oslo

Abstract

We owe to Ralph Lilley Turner the correct classification of Romani as originating from a central or inner form of Indo-Aryan. Turner also clarified that the “Dardic” elements in Romani have been borrowed into early Romani after its speakers had left their original home and reached the north-west of South Asia where they stayed for several hundred years before finally leaving the subcontinent. Until now, the extent of the “Dardic” influence on early Romani was poorly understood. In the present article much data has been put together which shows that this impact indeed is considerable. But it is intelligible only if we accept Turner’s hypothesis of a long stopover in north-western South Asia. The data presented below will also show that the notion of “Dardic” is too narrow in this context: the impact on early Romani, in fact, comprises linguistic elements and features found in Nuristani, Dardic and West Pahārī.

Keywords: Romani history, Nuristani, Dardic, West Pahārī, Indo-Aryan, linguistic borrowing.

1 Introduction

I have been working for some time on languages within the three Indo-Aryan subgroups¹ Nuristani, Dardic and West Pahārī.² They are spoken in the mountainous tracts between the Yamuna valley in the south-east and eastern Afghanistan in the north-west. Recently I have also been working with speakers of Norwegian Romani in a project financed by the Norwegian Research Council—to whom I want to express here my gratitude. It lent itself for me to combine the two fields of work and reconsider the question of the impact of languages of north-western South Asia on early Romani.³ It has been known for a long time that some influence does exist, but no detailed study has been done. The present article attempts to do this and I will show that not only Dardic but also Nuristani and West Pahārī need to be kept in view. My conclusions will be that their impact on early Romani is much more extensive than previously assumed.

The article begins with a succinct reevaluation of some questions concerning the early history of Indo-Aryan. It then recapitulates the most important arguments for the most likely geographical area from where the speakers of Romani must have originated. It will be seen that I agree to a large extent with the thesis suggested by R. L. Turner (1927), namely that Romani originally belongs to Central Indo-Aryan, respectively to the so-

1 The structuring of these subgroups is discussed in section 2.

2 The presentation of the language names follows the system found in R. L. Turner's *A Comparative Dictionary of the Indo-Aryan Languages*. There are, however, additional languages not found in Turner for which I have used simplified transliterations. For the varieties of Romani I have used own abbreviations. All language names, their linguistic affiliations and the abbreviations are found at the end of the article before the bibliographical references. Note also that if a word is said by me to be found, e.g. in Norwegian Romani this does not exclude the possibility that it may also be found in other varieties.

3 Throughout this article I use the notion '(languages of the) north-west' for the geographical area of the mountainous tract between the Yamuna valley in the south-east and eastern Afghanistan in the north-west. This covers the linguistic area of West Pahārī, Dardic and Nuristani. The notion 'west' I use for the catchment area of the Indus River.

called inner branch of Indo-Aryan (see section 3). By clarifying that Romani originated in Central Indo-Aryan and not in the area of the Dard languages in north-western South Asia, Turner rectified previous views held for instance by F. Miklosich, R. Pischel, and G. A. Grierson (see literature). I will support Turner's view with much data which prove a considerable influence on early Romani by Indo-Aryan languages which are found today in the mountains between the valley of the Yamuna and eastern Afghanistan.

2 Early Indo-Aryan

Instead of summarising the complex debate on the migration of speakers of Indo-Aryan into South Asia (see for instance Bronkhorst and Deshpande, Masica, Erdörsy, Renfrew, Marcantonio) I want to point out here that the whole discussion (including by authors who reject such an immigration) usually stops short with the beginning of Old Indo-Aryan in north-western South Asia. This gives the impression as if the further development of Indo-Aryan would not be of interest for this debate. But this is not the case. In a nutshell: (a) the fact that the most archaic forms of Indo-Aryan are found in north-western South Asia, and (b) the fact that it is possible to distinguish branches of Indo-Aryan whose vanishing point directs exactly to that north-western area are additional arguments that make it impossible to assume that the original home of Indo-Aryan is located elsewhere in South Asia.

The debate on inner and outer languages

Mainly in the fifth and sixth chapters of *Linguistic Archaeology in South Asia* the author Franklin C. Southworth discusses a controversial hypothesis of George Abraham Grierson. Grierson had suggested dividing New Indo-Aryan into three subgroupings which he called midland languages, intermediate languages and outer languages (Southworth 2005: 130). According to this model, West Pahārī would be an inner (or midland) and Kashmiri,

generally classified as Dardic, would be an outer language (see Masica 1991: 451 for a diagram). Grierson's hypothesis was not widely accepted because it was argued that there is too much diffusion and overlaying between the different IA languages so that no clear picture can emerge.⁴ It would be too long-winded to present here all the details of how Southworth not only defends Grierson's hypothesis but, in fact, modifies and places it on safer ground with the help of additional data. The main arguments for distinguishing between inner and outer languages (plus a transitional zone) are the following. All or most of the outer languages have (in the transitional zone the evidence is ambiguous) the following features which are missing with the inner languages (only the most relevant are quoted here): past forms in *-l-*; gerundive, nominal and future forms based on OIA *-(i)tavya-*; particular behaviours of the OIA vowels *r̥*, *i* and *u*; lexical evidence. Discussing questions of early dialect variations Southworth says (2005: 155): "In interpreting earlier evidence it is important to note that some scholars, for instance Chatterji, have assumed (tacitly or explicitly) that Pali and the Prakrits represent a stage intermediate between the earliest Indo-Aryan and the modern spoken languages. Others take the position that, from the Vedic period onwards, there were varieties of Indo-Aryan which were outside the "high" tradition ... If this was true in Vedic times, it would have been even more true during the MIA period when the Indo-Aryan languages were spread over a much larger territory. Thus it is reasonable to assume that along with the attested literary Prakrits there were also "colloquial Prakrits" which never appeared in writing."

Relevant here is that with regard to the Ashokan dialects, Southworth arrives at the following conclusion (2005: 167f.): "Bloch's three-way division of the Ashokan dialects (Center-East, Northwest, West) can be resolved into an earlier two-way division between the Northwest and the remaining dialects." North-west is reflected in the inscriptions of Shahbazgarhi and Mansehra, and the remaining dialects are classified by Southworth as West (Girnar, Sopara), Midland (Kalsi) and East (Dhauili, Jaugada). It is

4 For instance Masica holds this view, see Masica 1991: 460.

a well-known fact that the Dard languages are modern descendants of Prakrits more or less close to the Ashokan Northwest dialect (that is, Gāndhārī). Southworth, however, does not include Dardic in his model. Instead, he gives several reasons in the fifth chapter of his book (p. 149, footnote 9) why he excludes the Dard languages from his further investigations: fragmentary knowledge about past forms in *-l-* in Dardic; inadequacy of the descriptive material; questionableness of Dardic being a genuine subgroup of Indo-Aryan. The author has already been criticized for this (see the review of Kulikov) and a look at his diagram on page 168 is revealing. The diagram illustrates, with the help of a series of isoglosses, the above-mentioned division between the north-west and the remaining dialects. So what he does here is to almost exclude from his modernized inner-outer languages model MIA Gāndhārī and NIA Dardic because they share so few isoglosses with the rest. On page 169f., Southworth summarizes: “Evidence for the existence of two distinct sociolinguistic regions, inner (North-Central) and outer (South-Eastern) Indo-Aryan was presented ... The totality of the evidence points to the existence of two sociolinguistic regions, each showing some internal uniformity *vis-à-vis* the other, which however were probably in at least intermittent contact throughout most of their history ... The evidence of Vedic dialects ... does not conflict with, and possibly supports, the inner-outer group hypothesis, in that the major dialect division in the late Vedic period is between a midland dialect and an eastern-southern dialect, with a transitional dialect in the area of Kosala, the modern Avadh—exactly where Grierson placed his intermediate group. The northern Panjab, less active at this time in terms of text production, forms a separate dialect area.” Southworth's goal is not just to present additional evidence for Grierson's original thesis but he also suggests (2005: 181ff.) “[a] reconstruction of the prehistory of outer Indo-Aryan.” Here follows a very short synopsis. On page 181 he says in the section on Indo-Aryan in the Indus valley: “By 1500 BCE, when the first hymns of the Rigveda are believed to have been composed, that portion of the Indo-Aryan speech community which was associated with the OIA texts was located in the upper Indus Valley ... Given the archaeological evidence for intrusive Central

Asian elements on the lower Indus ... it can only be assumed that OIA speakers also occupied this area by the end of the second millennium BCE ... as the OIA 'mainstream' society expanded eastward across the Indo-Gangetic divide ... its counterpart in Sindh probably did the same, following the route mentioned here, leading to Malwa, Gujarat, and the Deccan." According to Southworth's model, the southern migration movement turned south and ultimately east towards the eastern limits of the subcontinent, and both (language) movements met and mixed in the transitional area of Kosala (Avadh).

A third branch in Indo-Aryan

In a forthcoming book I will demonstrate in great detail the existence of a third branch of Indo-Aryan (besides the inner and the outer branch) which I call the north-western branch. It includes Nuristani, Dardic and West Pahārī. The inclusion of Nuristani seems to contradict the widely shared opinion that Nuristani constitutes a separate branch within Indo-Iranian. It is certainly true that only Nuristani reflects a pre-Old Indo-Aryan stage; however, Nuristani has to be included in the north-western branch due to the overwhelming number of features it shares with Dardic and West Pahārī in grammar, vocabulary and a common cultural heritage of which traces can be found at different places. Nuristani branched off at the time of Proto-Aryan, but its present geographical adjacency to Dardic must be very old. In fact, I believe that Nuristani never got spatially much separated from Dardic. If I may employ here an image: Nuristani, Dardic and West Pahārī are like three siblings, Nuristani being the eldest, Dardic the middle and West Pahārī perhaps the youngest. This image works reasonably well only with regard to the preservation of archaisms. But to be regarded as a separate north-western branch this is not sufficient. It has to be shown that these three language groups also share innovations which are not found in the other Indo-Aryan languages. I do this in my forthcoming monograph; here it has to suffice to just list some of these innovations. The most important are: (a) preservation of a three- or

two-stepped system of sibilants (e.g. *s*, *š*, [*ʃ*]) which facilitated the *innovation* of a three- or two-stepped system of affricates (e.g. *dz*, *dž*, [*dʒ*]) as a result of depalatalization which had an impact on all three language groups over a long period of many centuries; (b) converbs (absolutives) and past forms built with an element *-t-*; (c) use of a non-aspirated auxiliary *tū* 'is; was' which historically derives <OIA *sthitá-* 'standing, settled'. In addition to these innovations there are a number of other innovations whose geographical extension within Nuristani, Dardic and West Pahārī is not as comprehensive as the three quoted elements; still they too are only found there. Pointing out that the three language groups share a common vocabulary is, at first sight, an *argumentum e silentium*; however, it is the sheer amount of shared vocabulary which is necessarily convincing. Romani participates in this, as I will show in this article, and there are even a few instances of the common north-western cultural heritage that come out in Roma traditions. Thus I argue that Romani belonged originally to the inner group of Indo-Aryan, but it has been influenced quite strongly by languages of the north-western group. This can only be explained by assuming that the speakers of early Romani stayed, after having left their original home, over a long time (perhaps several centuries) in the north-west of South Asia.

3 Romani

Romani belongs originally neither to the outer nor to the north-western branch of Indo-Aryan. This means that the speakers of Romani originated from the area of the inner branch. Their original home could have been somewhere in the area where today Hindi is spoken.⁵ In order to substantiate this I summarize here

5 In a paper from 1919 Grierson gave up his former idea that Romani is related to the language of the Doms of Bihar, however, only to make another untenable claim, namely to locate their home in the Dard languages area (see Turner 1927: 4).

Turner's most important arguments (1927) for allocating Proto-Romani to the inner group:⁶

Early innovations

1. Syllabic OIA *r* got changed at an early stage into *a*, *i*, or *u*; however, there are geographical differences: it changed into *a* in the south-west and south, elsewhere into *i* or *u*, and in the north-west it was partially preserved as *ri*. In Romani, *r* got changed into *i* or *u*: *kišlo* 'thin' < OIA *kṛśá-* 'lean, thin', *bukko* 'intestines' < OIA *vṛkká-* 'kidneys'. Turner says (1927: 8) that Romani *ric* 'bear' (OIA *ṛkṣa-* 'bear') is, like Hindi *rīch*, a loan from a Hill language. However, very similar forms are found in many other modern Indo-Aryan languages, and there is Prakrit *riccha-*. It is therefore unclear when and where the speakers of early Romani borrowed this word.
 2. The consonant group OIA *ṛt* led to retroflexion both in the east and in the north-west, whereas the dental was preserved in the central and south-western languages (although there are exceptions like Hindi and Panjabi *maṛak* 'plague' < OIA *mṛtakka-*). Romani agrees with the central and south-western languages: *mulo* 'dead' < OIA *mṛtá-* 'dead' where the *-l-* goes back to older *-t-* and not *-ṛ-* (which would have resulted in *-r-*).
 3. The OIA sound *kṣ* developed either into *ch* (*çh*) or *kh*, but the geographical picture is quite confusing. In the north-west there is a strong tendency for *ch* (*çh*), but both Hindi and Romani usually show *kh*: Romani *jakh* and Hindi *ākh* both 'eye' < OIA *ākṣi-* 'eye'. Turner lists four Romani words where OIA *kṣ* appears as *č(h)*, but it also does so in central languages like Hindi: *rič* 'bear' (see above 1.), *čhar* 'glowing ashes' < OIA *kṣārā-* 'corrosive' (Hindi *chār* 'alkali, ashes'), *čhuri* 'knife' < OIA *kṣurá-* 'razor' (Hindi *churā* 'dagger, razor'), *čulo* 'a little' < OIA *kṣulla-* 'small' (Hindi *chullū* 'childish').
- 6 Under every topic Turner also discusses various problematic cases. Only some of them are taken up here in case they are seen as relevant for my arguments.

4. The OIA consonant groups *śm*, *ṣm*, *sm* later on developed into *sp*, *ss* (*śś*), *mh*, *pph*, *mbh*. Again it is difficult to draw a clear picture. The change to *sp* and *ss* (*śś*) is largely limited to the north-west (but see the wide dispersal of OIA *raśmī*- ‘rope’ as *raśśi*, *rassi* etc.); there is only little evidence for *pph* and *mbh*; the most frequent and widespread is *mh* which also applies for Romani: *ame* ‘we’ < OIA **asme*. Also here Romani agrees with the central languages like Hindi, but disagrees with the north-western languages including Sindhi, Lahnda and Panjabi.
5. For the development of the OIA consonant groups *tv*, *dv*, *tm* there are only very few cases. Thus it suffices to say that for OIA *tm* Romani has *p* as in *po* or *pes* (oblique) ‘self’ < OIA *tmán*- ‘one’s own person’. According to Turner (1927: 14) here Romani differs from Dardic. However, there is Dardic Torwali *pāe*, Kashmiri *pān* ‘self’, West Pahārī Pōgulī *panun* ‘own’ and Sindhi *pāṇa* ‘reflexive pronoun’. So the Romani word *po* is probably a loan word from the north-west.
6. The change of OIA initial *y*- to *j*- is very widespread in Indo-Aryan but has not (completely) taken place in Dardic, Sindhi and Sinhalese. Romani belongs to the majority group: *džov* ‘oats’ < OIA *yava*- ‘barley’.
7. Intervocalic OIA *-m*- has been preserved in Dardic, Sinhalese and, to a certain extent, in West Pahārī. Otherwise it changed into a nasalized vowel and in Romani the nasalisation got subsequently lost: *kovlo* ‘soft’ < OIA *komalá*- ‘tender, soft’.

Conservations

1. Turner states (1927: 17): “Romani preserves *-t*-, probably *-d*-, perhaps *-th*-, and less certainly *-dh*-, under the form *l* in the European and Armenian dialects and *r* in the Syrian.” For instance *gili* ‘song’ < OIA *gīti*- ‘singing’, *len* ‘river’ < OIA *nadī* ‘river’ (with metathesis).
2. The three OIA sibilants *ś*, *ṣ*, *s* are preserved as two in European and Syrian Romani, but reduced to one in Armenian Romani. In European Romani *ś*, *ṣ* have merged into *š*: *šošoj*

- 'hare' < OIA *śaśá-* 'hare'; *šov* 'six' < OIA *śaś-*, nominative *śaṭ* 'six'; *sap* 'snake' < OIA *sarpá-* 'snake'. Among the modern Indo-Aryan languages, three sibilants have been preserved in Dardic (and Nuristani) and two (*ś*, *s*) in West Pahārī (and Ḍumākī⁷); in all other languages they got reduced to one.
3. OIA labial or dental + *r* are usually preserved in Romani as well as to a varying extent in the west and north-west as in Sindhi, Lahnda, Dardic and West Pahārī. Romani *trin* 'three' < OIA *trīṇi* 'three'; *prasal* 'to mock at, laugh at' < OIA *prahasati* 'bursts into laughter'. Velar + *r* have not been preserved in Romani, but it has been so occasionally in Dardic and West Pahārī.
 4. The OIA clusters *ṣṭ(h)* and *st(h)* have been preserved in European and Syrian Romani and, to some extent, in Dardic (and Nuristani) and West Pahārī. Romani *vušt* 'lip' < OIA *óṣṭa-* 'lip'; *vast* 'hand' < OIA *hásta-* 'hand'.

Turner concludes from these observations (1927: 22f.) that at the time of the Ashoka inscriptions the clusters with sibilants had already been changed except in Girnar (south-west) and Shahbazgarhi (north-west). Since Romani cannot be associated with either, he concludes—in my eyes absolutely correctly—that the speakers of Romani must have left their original home already before the time of the Ashoka inscriptions. He further observes that at about 250 AD *-d-* and perhaps *-t-* still turned up in Kharoṣṭhī documents of Khotan where also the clusters with sibilants and clusters with stops and *-r-* survived. Then he concludes that the speakers of Romani, which possessed the above features at the time of the departure from their original home, could preserve them in the north-west (where they stayed a considerable time) when the features disappeared in their original home.

7 Ḍumākī has also a retroflex *ṣ*, but this is only found in borrowings from surrounding Dardic and Burushaski.

Later innovations

1. The first innovation here Turner discusses is this (1927: 24): “A breathed consonant preceded by a nasal has been voiced in Sindhi, Lahnda, Panjabi, the whole Dard group (except perhaps Gawar-Bati; and not in Kafiri), and all the Pahari dialects as far as, and including, Nepāli (except for a few small enclaves). This is the normal treatment of Romani.” For example: *dand* ‘tooth’ < OIA *dānta-* ‘tooth’, *bango* ‘crooked’ < OIA *vañka-* *‘bent, crooked’. However, Masica (1991: 203) points out that there are several West Pahārī languages where this process has not taken place, and for instance in Bangani there are a number of doublets with unvoiced and voiced stops. We are dealing here with an incomplete phonological process. According to Turner, this sound change is not found in the Ashoka inscriptions but only by the time of the Kharoṣṭhī documents. He takes this as a possible hint that the speakers of Romani had reached the north-west of South Asia before Ashoka, because then “... it is not surprising that they should have shared subsequent innovations of that linguistic area” (1927: 24). Then follows the first example for this claim:
2. Metathesis of *r*. Turner states (loc. cit.): “In Sindhi, Lahnda, Dardic, and West Pahari, when the group *r* + consonant or consonant + *r* occurs in the middle of a word, the *r* is transposed (after the accompanying consonant has been doubled) and pronounced after the first consonant of the word. Thus Si. *ḍrigho* ‘tall’ (*ḍirghāḥ* > *ḍirgghāḥ* > *drīgha*) ...”⁸ An example from Romani: *tradel* ‘to drag, drive away’ < OIA *tārdati* ‘sets free’. Another example, where there is, however, metathesis without a “supporting” consonant, is *brivel* ‘to comb wool’ < OIA **vivarati* ‘uncovers’.⁹
3. Initial OIA *v-* has become *b-* in the central and eastern groups as well as in Pahārī, Ḍogrī and many Dard languages probably

8 I think that Turner is here not completely right. At least sometimes this process went like this: *ḍirghāḥ* > **ḍrīghāḥ* > *drīgha* (see Zoller forthcoming).

9 Many more examples, also involving borrowed words, are found in Boretzky 2005.

at a relatively late time. The original approximant has been preserved in the west and various parts of the north-west. This inconsistent picture seems to be reflected in the fact that we find *b-* in European Romani but *v-* in Syrian and Armenian Romani: European *berš* but Syrian *vars* both ‘year’ < OIA *varṣá-* ‘rain’.

4. The Turner article contains a small section on morphology. An interesting point made here by him is the fact that in European Romani the nominative singular masculine ends in *-o* because this may be a hint that Romani did not belong to the Māgadhi area where the ending is *-e*.

Summarizing the above points, Turner concludes (1927: 31) that there was “... an original connection with the Central group, and a subsequent migration to the North-west group.”¹⁰ Then Turner continues to substantiate his claim with regard to vocabulary where he demonstrates that the core vocabulary of Romani belongs to the central group. He continues on p. 32: “... it would not be surprising to find that the Gypsies had borrowed some words from the North-western languages, among which they must have lived for several centuries after leaving the Central group.” He shows that the words for ‘four’ and ‘six’ come from the north-west (the former seems to be connected with Nuristani). The few examples given by Turner will be supplemented in this article with many more words which further corroborate his view.

4 Words already known (or suspected) to be of IA origin and which (may) have been borrowed from the North-western branch

Here the main sources of information are Turner’s *Comparative Dictionary of the Indo-Aryan Languages*, Norbert Boretzky and Birgit Igla’s *Wörterbuch Romani – Deutsch – Englisch für den*

10 Note that Turner is not concerned here with the issue of inner and outer branch. However, his notion of Central group resembles that of the inner branch.

südosteuropäischen Raum, and Mathias Metzger's 'Etymological glossary of Indic words in Romani' (besides the sources quoted in the Literature section). Note that in the following section the known borrowed words are not just listed, but usually supplemented with additional information. It will also be seen that it is not always clear whether a Romani word is inherited or borrowed. Concerning the above-mentioned sound shift of dental stops to *-l-* both in Romani and in Dardic I assume that Romani words displaying it are all borrowings from the north-west because: (a) there is no evidence for a parallel development to this in the inner languages; (b) European Romani shows parallels with Dardic regarding this sound shift in case of *medial* dental stops (initial stops were not affected in European Romani and Dardic); (c) Armenian Romani shows parallels with some varieties of Nuristani and some East Iranian languages because initial dental stops also underwent this change there. The following words are ordered according to the North Indian alphabets.¹¹

Rom. *arakhel* 'to protect; to find'

Compare OIA *āraṅṣati* 'watches over, defends' (1298).¹² Modern descendants of the verb are only found in Waigali (*araç-* 'schützen') and Sinhalese; and of the noun OIA *āraṅṣa-* 'protection' (1297) again only in Waigali and Sinhalese. Since, however, the Romani word has *kh* for OIA *kṣ* and not an affricate as Waigali, it is not clear whether it is inherited or borrowed.

Rom. *avdive(s)* 'today; in the course of time'

This is a compound connected with OIA **ā nūnam* 'up to now' (1180) + *divasá-* 'day' (6333). Sub 1180 see Ashkun *yanũ* and Waigali *önũ* both 'today'. Forms without prefixed *nūnam* 'now'

11 I have cautiously standardized the many different transliterations and transcriptions used in the various quoted sources. Care should be taken that in Romani transcription the sound *j* is an approximant, whereas the same letter is a palatal voiced affricate in languages spoken in South Asia; the Romani sibilant *š* is equal to the Indo-Aryan sibilant *ś*, Romani *c* corresponds to Indo-Aryan *ç*, whereas *č* corresponds to *c* and *ć*.

12 Figures in parentheses refer to entries in Turner's *Comparative Dictionary of the Indo-Aryan Languages*.

are more numerous, but still also limited to Nuristani and Dardic. Synonym compounds are again found only in the north-west: Pašāi *nun-diwoś*, Tregami *nū-wās* ‘today’ (both lit. ‘now-day’), Torwali *až-dī* ‘today’ (lit. ‘today-day’); cf. also Kalasha *aǰ ádua* ‘today’ (lit. ‘today noon’) with the second word < OIA *ardhadivasa-* ‘noon’ [654]). The Romani word is thus most likely a borrowing from the north-west.

Rom. *ašarel*, *ušarel*, *šarel* ‘to praise’

According to Turner, Boretzky and Igla perhaps < OIA *ślāghate* ‘values, praises’ (12734) (plus a prefix). Compare e.g. Sindhi *sarāhaṇu*, Panjabi *sarāhuṇā* etc., all ‘to praise’. Turner explains the change *l* > *r* due to the spread of western IA forms where such changes do occur. An alternative derivation < OIA *úccarati* ‘rises, utters, speaks’ (1641) is semantically problematic. A derivation < OIA *śāthāyate* ‘flatters, beguiles’ or *śālate* ‘praises’ (suggested by Endre Tálos) faces the difficulty that the words are not found in documented languages. Compare, however, Bangani *śoreṇɔ* ‘to praise s.o.’ with palatal sibilant. Connection with OIA *ślāghate* is therefore most likely, which means that the word has entered early Romani in the west or north-west.

Rom. *ukljel* ‘to climb, ascend’

Compare OIA **utkalati* ‘goes out or up’, *utkalita-* ‘rising, prosperous’, *utkalyate*, *utkālayati* ‘drives out’ (1716). Modern descendants are found in many Indo-Aryan languages but not in Nuristani, Dardic and Sinhalese. However, the meaning ‘to climb, ascend’ is only shared with Pahārī languages: West Pahārī Koṭgarhī *ukəlno* ‘to climb, ascend’ and Bangani *ukləṇɔ* ‘to mount (as a bull a cow)’, Garhwālī *ukaḷnu* ‘to climb, ride’ and Nepālī *uklanu* ‘to ascend’. Still it cannot be said with certainty whether the Romani word is inherited or a borrowing.

Rom.G. and Rom.Germ. *ulo* ‘born’

The word derives < OIA *bhūtá-* ‘become, been, past’ (9552); cf. also Rom.H. *uló* ‘was’ and other Romani varieties *ulō* ‘(he) became’. Besides the *-l-*, which reflects OIA *-t-* and which is thus an indication for borrowing from the north-west, it is also the

semantics which points into the same direction. There is Bangani *ɔnɔ* ‘to be born’ and *uaɔnɔ* ‘to give birth’, Deogari *uaɔnɔ* ‘to give birth’ and Khas̥dhari *hui* ‘delivered (said about a baby)’. Even though these three languages are spoken in the same area and I am not aware of further semantic parallels in the north-west, it is quite likely that the Romani words are borrowings from the north-west.

Rom. *kiral* ‘cheese’

See OIA *kilāṭa-* ‘inspissated milk’ (3181); modern descendants are only found in Nuristani and Dardic; the Romani form is thus a borrowing (it underwent metathesis in the second syllable).

Rom. *kišaj* ‘sand’

Whether we are dealing here with an Indo-Iranian word (Lubotzky p. 5) or a Wanderwort (Cheung 2002: 227) is not quite clear. But the Romani word is certainly not a direct derivation < OIA *sikatā-* ‘grain of sand; sand, gravel’ (13386) but either a north-western or Iranian borrowing (with syllable metathesis). The lemma is found between West Pahārī in the south-east and Kurdish in the northwest, and there might be some distant parallels in Central and South India. There must have been considerable (re-)borrowings in the north-west, probably involving Iranian languages. In NIA the word is attested in Dardic, and West Pahārī, including Nuristani: Kati, Waigali, Ashkun, Dameli, Khowar, Kalasha, Pašāī, Shumashti, Gawar-Bati, Woṭapūrī, Bashkarīk, Savi, Phalūrā, Indus Kohistani, Kashmiri, and some West Pahārī varieties; e.g.: Pašāī *sēo*, Shumashti *sīū*, Woṭapūrī *sigit*, Bashkarīk *sigit*, Kashmiri *sekh* ‘sand, file’, Khashi *sikk* ‘gravel’, Bhalesi *sikkā*. Other Nuristani and Dardic languages have *ś-/ć-: Kati *čü*, *čüyü* ‘sand’, Waigali *šo*, Khowar *šugūr*, Kalasha *šīgou* and *šīgol-*. Iranian: Ossetic *sygyt* ‘earth(-matter)’, Sogdian *šykth*, Pashto *šəga* and Kurdish *sigit*. Lubotzky (p. 10) points also to Kannada *usiku*, *usigu* ‘sand’ and there may be a distant connection with Munda Santālī and Muṇḍārī *gitil* ‘sand’.

Rom. *koř* ‘neck’

The word derives < OIA *kroḍa-* ‘breast, bosom’ (3607) but has semantic parallels with the meaning ‘neck’ in modern languages

only in Dardic and Kashmiri: Kalasha *kruṛa*, Phalūrā *kirōṛ*, Kashmiri *koru* all ‘neck’. This semantic change is apparently due to overlap with phonologically similar derivations < OIA *kṛkāṭikā* ‘joint of neck’ (3419) which are limited to Nuristani, Dardic and West Pahārī. Thus the Romani word is a borrowing from the north-west. Note, however, that Rom.T. has, besides *korri* ‘neck’, also the word *kirko* ‘throat’. This resembles e.g. Bangani *kerko* ‘neck’ and Waigali *kir’ik* ‘Nacken’ (both derive < 3419) so that the Rom.T. word may have changed its original meaning through the influence of *korri*. Thus, also this word is a borrowing from the north-west.

Rom. *khil* ‘butter’

The word derives < OIA *ghṛtā*- ‘ghee’ (4501), see Indus Kohistani *ghīl* ‘ghee’. Again this word must be a borrowing from the north-west.

Rom. *xandžuvalo* ‘miserly, greedy’

Compare OIA **kacca*- ‘raw, unripe’ (2613) and the following Dard forms with aspirated initial consonant: Shina *khāčār* ‘ingratitude’, *khāčēlū* ‘miserly’, Dameli *khača* ‘dirty, bad’, Kalasha *khāčā* ‘bad, dirty (?)’, Phalūrā *khāču*, *khāčulo* ‘dirty, bad’. A borrowing of the Romani word from the north-west is very likely.

Rom. *xanřudel* ‘to scratch, scrape’

The word belongs to OIA *kaṇḍū*- ‘itching, the itch’ (2688) + the verb *del*.¹³ The derived OIA lemma *kaṇḍūyāti* ‘scratches’ (2689) has several modern descendants with initial aspirated stops in the west and north-west: Shina of Gilgit *khaṇōīkī*, Sindhi *khanhaṇu* ‘to scratch’, *khanvaṇi* ‘scratching’, Panjabi *khanūhṇā* ‘to itch’, the Kocī variety of West Pahārī *khəṇamiṇe* ‘itching’, *khəṇamiṇo* ‘to itch’, the Kotgarhī variety of West Pahārī *khəṇeuṇo*. According to Boretzky and Igla, the word was perhaps also influenced by Persian *xeneš* ‘Jucken’ and perhaps even by Armenian *xandz-el* ‘versengen; anbrennen’. And according to Turner (*Addenda and*

13 The word with this verbal ending is again discussed below in section 16.

Corrigenda), the aspirated forms of the IA words are influenced by OIA *kharju*-¹ ‘itching, scratching, scab’ (3827) or *khára* ‘hard, sharp, pungent’ (3819). This, together with the fact that the lemma is not found in inner NIA, suggests that the Romani word is a borrowing from the north-west or west, but got perhaps also influenced by words from other languages.

Rom. *gili* ‘song’

See Indus Kohistani *gīl* ‘song’, Šāṭōṭī *gīl*ⁱ and Savi *gīli* ‘song’ which are < OIA *gīti*- ‘singing’ (4168). According to Turner the Dard forms have an extension *-l-*, but this is untenable because of the final high front vowel in Šāṭōṭī and Savi, and because it would mean that the considerable number of words e.g. found in Indus Kohistani with *-l-* going back to a dental stop would all be borrowings from an unknown Dard language. Turner quotes also Rom.Syr. *gref* and asks “whence *f*?” Compare this, however, with Rom.N. *jilipa* ‘song’ and Rom.S. *gijepa* ‘song’ both of which employ the well-known suffix *-(i)pa*.

In connection with ‘song’ there are also interesting words for ‘musician’: The common Romani term is *bašaldo*,¹⁴ but there is also Rom.T. *bašado* ‘violinist’ and Rom.S. *bāši-moš* ‘gårdsmusikant’. Rom.S. has in addition a word *ghildo*, translated as ‘fest, party’. This, however, cannot be the original meaning which must have been ‘musician’. Compare the designation *Ghilabari* for a Roma group living in Romania who are professional musicians (Berger 1985: 779); their name corresponds to that of the South European *Gitanes*, which also means ‘musicians’.

The Rom.S. word *ghildo* has an exact correspondence in Indus Kohistani *gildō* ‘singer’. The aspiration in *ghildo* and in *Ghilabari* is spontaneous (more on this see below in section 11) and has correspondences in West Pahārī Chinali *ghit* ‘song’ and *ghitāru* ‘singer’, and in Paṅgwālī *ghūt* ‘song’. The first element of *bašaldo* is related to *bašalel* ‘to play (an instrument)’ which in turn is related to *bašel* ‘to bark, roar, howl’ which derives < OIA *vāśyate* ‘roars, howls, bellows, lows, bleats, sings (of birds)’ (11589). The

14 Discussed again below in section 15.

word *bašalel* probably contains an *-l-* transitive/causative suffix. Modern descendants of OIA *vāśyate* are limited to Nuristani, Dardic, West Pahārī, Kumaoni, Nepālī and Gujarati; it is thus not a central language word.

The words *bašaldo*, *ghildo* and *gildō* are compounds with a second element *-do*. This element derives < OIA *dādhāti* ‘places, lays on, gives, seizes’ (6145), but regarding the exact meaning one needs to consider the meanings given for the precursor PIE **dheh₁-*, namely ‘to put, lay down, sit down, produce, make, speak, say, bring back’. The OIA verb has (almost) exclusively survived in Nuristani and Dardic whereas in the other Indo-Aryan languages it was displaced by the very similar OIA *dādāti* ‘gives’ (6141). That we are indeed dealing with OIA *dādhāti* ‘places’ is corroborated by Indus Kohistani *gīlmār* ‘singer’ the second component of which comes from *mār.áṽ* ‘to kill (< OIA *māráyati*¹ ‘kills’ [10066]) which also means, e.g. in Hindi *mārnā* ‘to perform an action with vigour’. Thus, the underlying meaning of *gildō*, *ghildo*, *gīlmār* is something like “one who produces/belts out/performs a song”. It seems that we are dealing here with a very old compound. And indeed, the compound has a striking parallel compound building in the Celtic word *bard* which goes back to PIE **g^wṛh₂-d^hh₁-o-* ‘praise-maker’ (West 2007: 27). Whereas the first components of the compounds obviously have several different words associated with ‘song’ (in case of the Celtic word it is ‘praise’, in case of the Indic words it is ‘sing’), it is remarkable that the second component in the PIE reconstruction is exactly the same verb which we identified above: PIE **dheh₁-* ‘to put, produce, speak etc.’. Of course I am not in a position to say whether we deal here with a common Indo-European heritage or whether this is a matter of two independent developments at the two ends of the Indo-European world.

Rom. *gelo* ‘went, gone’

This is the preterite of *džal* ‘to go’ and derives < OIA *gatá-* ‘gone’ (4008); the word is a borrowing because of *-t-* > *-l-*.

Rom. *gošni* ‘cowdung’

The word is a compound going back to OIA *gó-* ‘cow’ (4255) plus **śakana-* ‘dung’ (12238). The latter is, according to the

information in Turner, limited to Nuristani and Dardic (but a sideform **chakana-* is found in Marāṭhī and Koṇkaṇī). There is also, but with a deviating semantics, Bangani *goṣṇi*¹⁵ ‘outdoor fireplace during monsoon in which dried cowdung is burnt in order to keep away biting flies’ which is an extension of *goṣu* ‘dried cowdung’ which itself derives < OIA *gośakṛt-* ‘cowdung’ (4333). This lemma is found in the north-west but also in Hindi. The likelihood that this is a north-western borrowing is perhaps enhanced by the fact that similar compounds tend to be typically found in the north-west. In any case, a direct parallel is Ashkun *ga-sā* ‘cowdung’ (with trace of a nasal consonant), and semantically comparable is Khashi *kuster* ‘dunghill’ which goes back to OIA **go-stara-* (see 13685) (devoicing of mediae is quite common in Khashi).

Rom. čiriklo ‘bird’

Compare OIA *caṭaka-* ‘sparrow’ (4571). This is a widespread lemma in Indo-Aryan. However, phonetically fairly close are Indus Kohistani *čaklú* ‘bird’, Kalasha *čilīngi* ‘sparrow’ and Savi *čunkerí* ‘Vogel’. Since there is also Romani *čirikli* ‘hen’ there might be Iranian interference.

Rom.Arm. čhen ‘female genital’ and Rom.Dol. čindi ‘vulva’

Finck (1907: 71) rightly suggests connection with OIA $\sqrt{CHID}$ ‘cut’. A comparable semantics (but no exact morphological correspondence) is found only in the north-west in Kotgarhī *chéuri*, Kocī *chéure*, Jaunsāri *chēori* all ‘woman’, Bangani *čevēr* ‘girl, woman’, Deogari *cheuri* ‘married woman’ all < OIA **chedu-* ‘cut, slit’ (5067b) with a *-ta-* extension. The Rom.Arm. and the Rom.Dol. forms are participles, compare Rom.Dol. *čindo* ‘geschnitten’. However, also related are widespread derivations < OIA **chinnāli-* ‘adulteress’ (5058). It is therefore difficult to say whether the word is inherited or borrowed.

15 The *-ṣ-* is an allophone of *-ś-*.

Rom. *čhela* ‘smallpox’

The word is a combination of elements of a derivation < *śītalā-* ‘(goddess of) smallpox’ (12490) and < *śīṭala-* ‘cold’ (12487).¹⁶ Only in Dardic languages has this lemma semantically split into ‘cold’ and ‘fever’, compare Gawar-Bati *śal* ‘fever’ but *śalá*, *śolś* ‘cold’, Savi *śal* ‘fever’ but *śalo* ‘cold’. In the other modern IA languages only the meaning ‘cold’ is found. The original meaning ‘cold(ness)’ is preserved in Romani *šil* which is < OIA *śīṭa-* ‘cold’ (12485). A parallel to the affricatized Romani word is found in Kalasha and Burushaski *čilá* ‘cold, cold season’. For the Romani word Boretzky and Iglá suggest irregular phonological development in a taboo word, but now we see that these irregularities are geographically located in the north-west from where this Romani word must have been borrowed. I am not aware of a *śītalā* cult beyond Kashmir, but ideas concerning smallpox—the goddess both sends out and heals the disease—may have been more widespread. When she disperses smallpox this is accompanied with fever; and if she heals the disease she does so with cold water.

Rom. Wel *džanel* ‘to bear (a child), be born’

The word derives < OIA **jāṇayati* ‘begets, bears’ (5192). Modern descendants are found only in Khowar and Sindhi. Therefore borrowing appears quite likely.

Rom. *džamutro* ‘son-in-law’

The word derives < OIA *jāmātr-* ‘daughter’s husband’ (5198). It looks like a north-western borrowing because Romani does not preserve medial *-m-*; however, also in some other languages which usually have lost OIA *-m-* the nasal consonant has been preserved as in Hindi *jamāī*. Thus, the exact origin of the Romani word is unclear.

16 On the closely related Rom.Syr. *sildā* ‘cold, unhappy’ see below in this section.

Rom. *thaj* ‘and, also’, Rom.T. *the* ‘and’

The word may derive < OIA *tathāpi* ‘even so, nevertheless’ (5647), but Turner is doubtful about the whole lemma. Still, what speaks in favour of a borrowing is first the fact that the meaning ‘and’ is limited to the north-west; second the Romani word possibly displays aspiration fronting (more on this below in section 10), which is a characteristic feature of Dardic and West Pahārī.

Rom. *dad* ‘father’

Compare OIA **dādda-* ‘father or other elderly relative’ (6261). The word might be a north-western loanword since the meaning ‘father’ is limited to that region.

Rom. *devel* ‘god’

The word derives < OIA *devatā-* ‘godhead, divinity’ (6530); it must be a borrowing because of *-t- > -l-*.

Rom. *nilaj* ‘summer’

Compare OIA *nidāghā-* ‘hot season’ (7193) and *nidāghakāla-* ‘heat of summer’ (7194). Turner is here slightly confusing as very similar Romani forms are quoted under the two lemmata. The first lemma is basically limited to Nuristani and Dardic, but it is also found in Oriya, however with a different meaning. I want to add here that a modern derivation of *nidāghakāla-* is also found in West Pahārī Khashī (or Khashali?) *nēlā* ‘summer’ (see Kaul 2006 I: 335). Thus the Romani word is most likely a borrowing from the north-west.

Rom. *porizen* ‘sieve’

Compare OIA **parivecana-* ‘sifting’ (7882) which has modern descendants in Dardic and Nuristani, but also close parallels in Eastern Iranian. Examples: Dameli *pareci* ‘sieve’, Kalasha of Rumbur *parec* (< OIA **parivecya-* ‘to be sifted’ [7882]), Ashkun *peica*, *peca*, Waigali *poca* (< OIA **pativecya-* ‘to be sifted’ [7730]). Compare also Pashto *pēzna* ‘sieve’ < **pativaičana*, and Shughni *parwej-* ‘to sow, sift’, *parwiz-* ‘to sift’, Parachi *paric-*, Roshani *parwizd*, Yazghulami *parwij* < **pariwaica*. Since the Romani word displays depalatalization it is certainly a loanword.

But it is unclear whether it was borrowed from north-west Aryan or from Iranian.

Rom. *phab*, *phabái* ‘apple’

The words derive < OIA **bhabbā*- ‘apple’ (9387). There are modern descendants only in Nuristani, Dardic, and Ḍumāki, thus this must be a loan word.

Rom. *phiko* ‘shoulder, shoulder-blade, support’

Compare OIA **sphiya*-, *sphya*- *‘scapula’ (13839). Parallels to the Romani form with -*k*- suffix and the meaning ‘shoulder-blade’ are again found only in the north-west: Kashmiri *phyoku* ‘shoulder-blade’, Shughni *fyak* ‘shoulder’ and *fiyak* ‘wooden shovel, shoulder-blade’, Ishkashmi *fayək* ‘shoulder’, etc. So the Romani word can either be a borrowing from Dardic or East Iranian.¹⁷

Rom. *phuḱol* ‘to swell; to blow’

Compare OIA **phūṭka*- ‘blowing’ (9102) with which the word is related, but not directly. Closer to the Romani form look Kalasha *phuṣ* ‘breath’, *phūṣik* ‘to blow (up a skin)’ and Indus Kohistani *phḷṣ*- *phḷṣ karāv* ‘to breathe’. The semantics suggest that there is a conflation of two OIA forms, namely *phūt*- ‘blowing puffing’ and √*ŚVI* ‘swell’, more exactly: *śvātrá*- ‘invigorating’ (but Mayrhofer suggests ‘Ausdehnung, Kraft; anschwellend, gedeihlich’). This word has been suggested to be the origin of the Kati and Prasun words for ‘rhubarb’. But here we can reconstruct from **phūt*- *śvātrá*- a proto-form **phūḱ* from which all quoted modern forms can be derived.

Rom. *phral* ‘brother’

See also Rom.N. *pral*, Rom.S. *phral*, *pral*, *prahl* ‘bror – brother’, etc. The forms belong to OIA *bhrātṛ*- ‘brother’ (9661). This is clearly a loan from Dardic even though there are no exact modern parallels, but compare Ashoka inscriptions from Mānsehrā and Shāhbāzgarhī *bhrat* ‘brother’ as well as Ḍumāki *birārā* and Khowar *brār* ‘brother’.

17 The word is discussed again below in section 8.

Rom. *bašel* ‘to bark, roar, howl’

The word derives < OIA *vāśyate* ‘roars, howls, bellows, lows, bleats, sings (of birds)’ (11589). Modern descendants of this lemma are only found in Nuristani, Dardic, West Pahārī and Kumaoni.

Rom.Eur. and Rom.Arm. *ma* negative of imperative ‘not’

The word derives < OIA *mā* ‘negative of prohibition (used with conjunctive and imperative)’ (9981). In NIA it is only attested in Nuristani, Dardic, and in Sindhi and Gujarati. Since it is not found unextended in any inner language, it seems likely that the Romani word is a borrowing.

Rom. *masek* ‘month, moon’

Compare OIA *māsa-* ‘moon, month’ (10104). Turner suggests the meaning ‘one month’ for the Romani word in order to explain the *-ek*. But this is implausible as there is also *mušteḱ* ‘palm of hand’ (see entry below in this section) with probably the same suffix. Parallels are found in Nuristani Prasun *masēḱ*, *mēsegē* ‘moon’ and Dardic Pašāī *mōyek* ‘moon, month’. Thus, the Romani word is most likely a borrowing from the north-west.

Rom. *murš* ‘man’

The word is a contamination of OIA *manuṣya-* ‘human; human being, man’ (9828) and *pūruṣa-* ‘man, male’ (8289). The closest parallel forms are found in the west in Sindhi *mursu* ‘man, husband’ and Khetrānī *murs*. Yet it is not quite clear whether the Romani word is a borrowing from the west.

Rom. *mulo* ‘dead’

The word derives < OIA *mṛtá-* ‘dead’ (10278). This must be a borrowing from the north-west even though there are no modern parallels. But the phonological development is the same as e.g. in *khil*.

Rom. *mušteḱ* ‘palm of hand’

Compare OIA *muṣṭi-* ‘clenched hand, fist’ (10221). The same word suffixed with *-k* is only found in Dardic Gawar-Bati *muṣṭāk*,

muṣṭīke ‘fist’ and in the Shina of Gures and Kohistan *mūṣṭāk* ‘fist’. The Romani word thus appears to be a borrowing from the north-west.

Rom. *mol* ‘wine’

The word may derive < OIA *mádhu-* ‘honey, mead’ (9784). A direct parallel is found in Prasun *mūlu* ‘wine’ and there is Burushaski *mel* ‘Wein (aus Trauben)’, but there is also Persian *mul* ‘wine’. So it is not clear whether the word was borrowed in the north-west or in Iran.

Rom. *lima* ‘mucus, phlegm’

The word derives < OIA *śleṣmán-* ‘mucus, phlegm’ (12744); similar looking derivations are only found in the west and north-west: Khetrānī *lim* ‘marrow’, Ḍumāki *līma* ‘mucus from nose’, Lahnda *lim* ‘phlegm, mucus from nose’, and West Pahārī Khashi and Bhadravāhī *limm* ‘mucus of nose’. Even though the phonological change *ṣm* > *m* can be inherited, the geographical limitation of *lima* forms to the west and north-west makes it likely that this lemma is a borrowing.

Rom. *lolo* ‘red’

Besides OIA *lōhita-* ‘red’ (11165) Turner also postulates OIA **lohila-* ‘red’ (11168), apparently on the basis of the occurrence of the lemma in various north-western languages to which he doesn’t want to assign a phonological rule involving a change of *-t-* > *-l-*: Waigali *lailāi-štā* ‘red’, Savi *lohīlō*, *loyol’o* ‘red’, Phalūrā *lohīlu*, *lāhoilo*, *lhōilo*, Chilis *līlo* ‘red’, Shina of Gilgit *lolyū* ‘red, bay (of horse or cow)’, Pašāi *lēle-šiōl* ‘fox’, etc. For Romani more likely is, however, borrowing of a modern descendant **lohila* that derived < the OIA lemma *lōhita-* in the north-west with typical change of *-t-* > *-l-*. Fussman (1972, entry 137) considers phonetic influence through derivations of OIA *nīla-* ‘dark blue’, but forms like Torwali *laūr*, Rom.Arm. *lohorī* and Rom.As. *lohri* all ‘red’ do not support this.

Rom. *šax* ‘cabbage’

Compare OIA *śāka-* ‘potherb, vegetable’ (12370). Modern descendants are found in Nuristani and Dardic, e.g. Waigali *čā* (without final consonant); Khowar *šax* ‘green vegetables’, Kalasha *šak* (with final consonant); Phalūrā *šo*, Shina *šā* and Indus Kohistani *šā* ‘vegetable’ (without final consonant). But the consonant is again preserved e.g. in Hindi *sāg* and Bangani *śāg* both ‘green vegetable’. The lemma looks to have been influenced by (repeated) *tatsama* borrowings, and thus it is not clear whether the Roma word is a borrowing or inherited.

Rom. *sastri*, Rom.G. *šastír* and Rom.T. *strast* all ‘iron’

The words derive ultimately < OIA *śastrá-* ‘instrument for cutting’ (12367). The Rom. and Rom.G. words could be inherited from a phonological point of view; however the fact that modern descendants are only found in Dardic, West Pahārī, Panjabi and Sinhalese makes it a strong borrowing candidate. Rom.T. *strast* ‘iron’ has direct parallels in the Pašāī dialect forms *leṣṭ* and *leiš* ‘knife’ which are, according to Turner, < older **straṣṭrī* < **straśṭrī* < **śtraṣṭrī*. This shows that the Rom., Rom.G., and Rom.T. forms are borrowings from the north-west, however from different places.

Rom. *šukar* ‘beautiful’

The word is usually derived < OIA *śukrá-* ‘bright’ (12506) but the typical NIA meaning is ‘bright, white, shining’. The meaning ‘beautiful’ appears to be limited to the north-west: Indus-Kohistani *šākər* ‘beautiful, pretty; lovely, charming (baby, young child)’, Burushaski *śákar* ‘lieb, geliebte(r)’ (but in case of the Burushaski word there is interference by the homophonous word meaning ‘sugar’); compare also Kalasha *š’ukri/šrúki* ‘naked (woman)’. Thus this word is perhaps a borrowing from the north-west.

Rom. *šut* ‘vinegar’

The word derives < OIA *śukta-* ‘become acid or sour’ (12504). Modern descendants are found only in Dardic: Pašāī *šut* ‘sour’, Khowar *šut* ‘sour’, *šutu* ‘buttermilk’, *šuti* ‘sourness’, Kashmiri *hoṭu* ‘decayed, tainted’, Indus Kohistani *šūt^h* ‘very sour; a vinegar made from apricots’. Also the related adjective Romani *šuklo*

‘sour’, which is < OIA **śuktala-*, has parallels only in Gawar-Bati *śutāla* ‘sour’ and Savi *śutāl*. Thus the Romani word is clearly a borrowing from the north-west.

Rom. *šerand* ‘pillow’

The word derives < OIA **śiraanta* ‘head-end’ (12448). It could be inherited from a phonological point of view; however, the fact that modern descendants are only found in Sindhi, Lahnda, Panjabi and West Pahārī makes also this word a strong borrowing candidate.

Rom. *šol/šil* ‘whistle’

Compare Bangani *śēr* ‘whistling’, Deogari *śērki* ‘whistling’, Indus Kohistani *šur šurī* ‘whistling’, Shina *šurūki d-* ‘pfeifen’ (which displays coronal consonant harmony), Kashmiri *śiriñ* ‘a whistle (formed with the lips), whistling’. These words cannot derive < OIA **śīṭṭa-* ‘whistle’ (13427) but require an allomorphic protoform **śuṭi-* ‘whistle’. From this all words here can be derived: Bangani, Deogari and Kashmiri display epenthesis (more on this below in section 13), but not Indus Kohistani and Romani. A change of *-ṭ-* > *-l* is rare in Romani, but it does occur; compare Rom. *džukel* ‘dog’ which is < OIA *jukuṭa-* ‘dog’. Boretzky and Igla also consider influence by Armenian *sul-em* ‘pfeifen’. Still, the Romani word is clearly a borrowing from the north-west.

Rom. *sasto* ‘healthy’

The word derives < OIA *svastha-* ‘well, healthy’ (13917). It could be inherited from a phonological point of view; however, the fact that modern descendants are only found in Dardic makes also this word a strong borrowing candidate.

Rom.Syr. *sildă* ‘cold, unhappy’

The word belongs to OIA *śītala-* (12487) ‘cold’.¹⁸ There are several Dard languages where the *-t-* has not disappeared: Paśāī *śidāl*, Shumashti *śidāl*, Torwali *śidul*, Phalūrā *śidālo*, Shina *śidālū*. They are regarded by Fussman (1972, entry 57) as semi-tatsamas (i.e. as later borrowings), but this is implausible. To be added to

18 The typical European Romani word is *šilaló* ‘icy, frozen’.

the Turner forms are Phalūrā *šid* ‘cold’ (noun), Savi *šid* ‘coldness; fever’ and *šideli* ‘cold’, but note that Phalūrā and Savi *šid* belong to OIA *śītā*-¹ ‘cold’(124859) (compare the discussion of Rom. *čhela* ‘smallpox’ above in this section). The word is thus a borrowing from the north-west.

Rom.G. *seli* ‘bran’

Compare OIA *satīnā*- ‘the pea *Pisum avense*’ (13116). The OIA lemma was so far not known to have a modern descendant besides Romani. But there are Bhadrawāhī *setu* and Bhalesi *sete* both meaning ‘bran’ (Kaul 2006 I: 327). Both words have preserved the original *-t-* (which is quite rare in West Pahārī, but there are other cases as well, and see the preceding entry). The background of this lemma is complicated by Rom. *šeli* (also Rom.T. *šelja*) ‘bran’ and Rudhari *šeli* ‘grain, bran (of maize)’ (Kaul: 2006 II: 274). Turner derives the Romani form < OIA *śadaka*- ‘unhusked corn’ (12287), but that cannot be the origin of the Rudhari word if one does not want to assume an isolated borrowing from an unknown Dard language since West Pahārī does not know a historical change *-t-*, *-d-* > *-l-*. On the other hand, the Romani word *šeli/šelja* does seem to have been borrowed from an unknown Dard language. So the matter is really unclear. Note, however, that the *-e-* in the words *seli*, *setu*, *sete*, *šeli*, *šeli/šelja* possibly resulted from the same phonological process, namely a so-called epenthesis (which is found in many areas where Dardic and West Pahārī is spoken; it is discussed below in section 13).

5 Words not yet known to be of IA origin and also borrowed from north-western languages

This section runs the risk of being characterized of using the *argumentum e silentium*. Indeed, this danger cannot be completely ruled out as it is always possible that new evidences for words at unexpected places come up. However, it is unlikely that *all* of the words presented below would finally be found to be known also in the inner group. This is simply also not possible because some of the words have undergone sound changes which definitely have

not occurred in the inner group. Also here I include words from all branches of Romani.

Rom.Arm. *akli* ‘a lie; untrue; insufficient’

Because of the preservation of the *-k-* I assume a borrowing from a north-western language even though the word is also found in the inner group. It derives < OIA *alīkā-* ‘unpleasing; untrue; a little’ (718).

Rom.Arm. *anlēš* ‘paradise’

There is no straightforward etymology for this word, but it may belong to OIA *anudeśa-* in the more literal sense of *‘adjoining land’. On the one hand, OIA *anudeśa-* has only the technical meaning ‘reference to something prior’. On the other hand there is, however, Shina *ōśō* (Turner: *ōśū*) ‘(male) guest’ which Turner derives < OIA **apadeśya-* ‘foreign’ (427) and which has been borrowed into Burushaski as *ośīn* ‘Besucher, Gast’. However, a derivation < OIA *anudeśin-* ‘residing at the same place’—but here with a suggested meaning *‘belonging to an adjoining area’—is more convincing also with regard to the initial nasalization of the Shina word (in the north-west *dēs* frequently means ‘village (and adjoining area)’). Admittedly, the palatal sibilant in *anlēš* cannot be the same as the OIA one and is probably an Armenian suffix (cf. Rom.Arm. *lehi*, *leji* ‘village’ and for possible suffixes Finck 1907: 50; and cf. Paśāī *dē*, Khowar *deh* and Rom.Syr. *dē* all ‘village’). I may also point out here that the word *deś* is (or rather was) used by the Prasun people in Nuristan in “*urdes*” ‘heaven’ and “*yurdes*” ‘paradise’ and ‘hell’ (Jettmar 1975: 51f.).

Rom.Burg. *eklik* ‘a little’

The word consists of two elements: *ek-* ‘a; one’ (< OIA **ēkka* ‘one’ [2462]), and regarding the second element compare Indus Kohistani *lāk^{hu}* ‘small; (a) little’, Iranian Saka *laka* ‘little, small amount’, Burushaski *luk* ‘a little’, Woṭa-pūrī *lukūr* ‘small’. In case of Bashkarīk *lūkuṭ* ‘shorter, younger’, Khashi *lokuch* ‘smaller’ and Kashmiri *lōkoṭ^u* ‘small, shorter, younger’ comparative suffixes have been added. According to Turner < OIA **lukka-* ‘defective’ (11072), but this is implausible because the geographical

distribution of this word with the meaning 'small' is limited to the north-west and to Saka. This makes it clear that it cannot belong to 11072, whatever its origin. Thus the Rom.Burg. word may be a borrowing from the north-west.

Rom. *kermuso* 'mouse, rat'

According to Turner, the first syllable is < OIA *ghara-* 'house'. But compare Pašāi *kavar-mūs* 'rat' and Persian *karmūš* 'muskrat'. It is unclear whether here also belongs Rom.T. *maskariš* 'mouse', but the forms suggest a compound word of unclear derivation but borrowed from somewhere in the north-west.

Rom.Arm. *konč, kõng, gūnč* 'beard'

The word goes back to OIA *gocchā-* 'furrow of upper lip' (4269) with modern descendants in Ashkun, Kati, Waigali, Gawar-Bati, Savi, Kashmiri, Bangani, Khaśdhari, Deogari and Bauri, all with the meaning 'moustache'. The Rom.Arm word is thus a borrowing from the north-west.

Rom. *čičalo₁* 'penis'

There may be an onomatopoetic dimension here, but there is a correspondence in Indus Kohistani *čičū* 'a small boy's penis'. Indus Kohistani *čičū* is different from *čič^h* 'nipple, breast' which is < OIA **cuccu-* 'female breast, nipple'. But there is perhaps either connection with Panjabi *cīcī* 'the little finger; the little toe' or the word is a north-western borrowing that goes back to OIA **śṛthilā-* 'loose, slack' (12601) (cf. Waigali *čičil'a* 'weich, leicht (facilis)', Prasun *čⁱčil* 'soft', etc.).

Rom. *čičalo₂* 'meat'

This word must have a different origin than the preceding one even though the two have been put together by Boretzky and Igla, and even though it also is of onomatopoetic character. It belongs to a fairly large group with examples known to me in the area between Ossetic and West Pahārī. Thus it seems to be a Wanderwort of unknown origin: Bangani *cīcaū* 'meat' (children's language), Deogari *cīcī* 'piece of meat' (children's language), Koṭgarhī *cī* 'meat, cooked meat', Indus Kohistani *čāčā* 'meat' (children's

language), Burushaski and Shina *ćáća* ‘Fleisch’ (Kindersprache), Ossetic *dzidza* ‘Fleisch’ (Babysprache).

Rom. *čhungar* ‘spit, saliva’

There may be a connection with Indus Kohistani *čurùk^h karāṽ* ‘to spit by pressing the saliva between the teeth in order to create a whizzing sound’, perhaps also *čarak^h karāṽ* ‘to squirt’, and Bangani *curuk-curuk* ‘sound of spitting and verbal *curkaṇə* ‘to spit’. But a derivation is unknown.

Rom. *čhomut/čhumut*¹⁹/*čhonut* ‘moon; moonlight’

The word is a compound with the first component deriving < OIA *jyótsnā-* ‘moonlight’ (5301) and the second < OIA *masta-*, **mastra-* ‘head; skull’ (9926). Compare the following words without derivations < OIA *jyótsnā-* but showing a homonymy of ‘moon’ and ‘head’: Kalasha *mastrúk* ‘moon, moonlight; month’, *másta* ‘brains’ and *mastrugón* or *mastrugonḍ* ‘scarecrow’ (with second component *-gon*, *-gonḍ* ‘stick’ < OIA *gaṇḍa-*² ‘trunk of tree from root to branches’ [3998] and thus basically meaning ‘stick with head’), Kati *m^rũšⁱte* ‘brain’ (< OIA **mastra-* with *r* fronting), Pašāi *māto* ‘moon’ and Pašāi dialect *mastrák* ‘Gehirn’. The semantic contamination was obviously caused by the similarity with derivations < OIA *māsa-* ‘moon; month’ (10104) as, e.g. in Ashkun *mās*, Niṅgalāmī *mas*, Gambīrī *mās* ‘moon’, Khowar *mas* ‘moon; month’ etc. A morphologically different but semantically identical compound formation ‘light-head’ exists in Pašāi *mōg-leš* ‘moon’ for which Fussman (1972, entry 84) considers combination of derivations < OIA *māsa-* and **locya-* ‘bright’ (11131), but it is phonologically better to derive the first component < OIA *mastīṣka-* ‘cranium’ (9926) which yields again a basic meaning ‘bright head’. Besides Romani, *jyótsnā-* as first and *masta-* as second component is found in Satlaj Group *jōth* ‘moon’, Inner Sirājī and Kuḷuī *dzōth*, Sainjī *dzōth* etc., and in *jōdhaiyā* ‘moon’ in the Lakhīmpurī dialect of Awadhi. The *-m-* of the Romani forms is of course not inherited but an allophone of the *-n-* of the first component (cf. Rom. *nilaj* ~ *milaj* ‘summer’), and the vowel *-u-* is

19 *čhumut* rarely also ‘beautiful girl’.

perhaps epenthetic reflex of an original form **mastu-* as it is found in OIA *mastulūṅa-* ‘brain’ (9926). Even though ‘moon’ as ‘bright head’ is found over a large area, the Romani forms are borrowings from the north-west because of the aspiration fronting (cf. MIA *jonhā-* ‘moon’ and see below section 10).

Rom.Zak. *labol* ‘to burn’ (itr.), Rom.T. *labarav* ‘to burn’ (tr.) and Rom.Lov. *lobo* ‘flame’

The lemma is found over a quite large area including West Pahārī, Dardic, Nuristani, Burushaski, some Iranian languages and Western Tibeto-Himalayan languages: Bangani *lɔpɔ* ‘to shine, sparkle (e.g., fire)’, *lapi* ‘flame; torch’ and *lupi* ‘flame; lamp; Kotgarhī and Kocī poetic *lupe* ‘flame’; Deogari *lɔp-lɔp* and Khasdhari *lɔp-lɔp* both ‘flaming, sparkling’; Kannaure *lɔpəŋg* ‘flame’; Chitkuli *lɔp-lɔp mē* ‘a flaming, sparkling fire’ (Zoller); Rudhari *lappi deni* ‘to fire, to heat by firing’ (Kaul 2006 II: 271), Gari *lɔp* ‘lightning’, Shina *lūpi-zhōiki-ī-zhēi* v.i. ‘burn’ and *lupōīkī* v.t. ‘burn (wood, etc.) light (fire lamp)’ (also noted as *lup-‘anzünden’*), Yasin Burushaski *lap* and *laláp* ‘shine, burn, light up; to beam’; Brokskad *lupras* ‘to burn, to kindle’ (with *-as* infinitive and perhaps with an *-r-* causative), *melɔp* ‘flame’ (a synonym compound with first component borrowed from Tibetan *me* ‘fire’) and *meleps* ‘fire fly, glow worm’ (like preceding but with (unclear?) extension) and probably *tralupis* ‘to shine’ (with unclear first component *tra-*); Indus Kohistani *lɔp-lɔp^h ho-* ‘to light up, shine, sparkle, glitter’, Waigali *luppa(h)* ‘lamp, torch’ and *lap’a* ‘Fackel’, Khowar *lápêik* ‘to glitter’ and Yidgha-Munjī *lapoir* ‘glitters’. The word seems to be of Proto-Indo European provenance and derive < PIE **lap-* ‘shine’ in which connection Mallory and Adams state (2006: 329) that this root “... may have been specifically related to the brightness of fire.”

Rom.Arm. *lorel* ‘to find, discover’

This word is somehow < OIA *loḍayati* ‘agitates’ or its sideform *lóḍati* (11080). Several of the modern descendants do not have the OIA meanings quoted here but meanings comparable with the Armenian Romani word: Panjabi *rolnā* ‘to sift (coarse from fine, rice from husk)’, Khowar *loḷik* ‘to look for’ and Bangani *lorɔ* ‘to

search for'. Bangani *lotnɔ* 'to rock, sway; to fall (down), collapse' shows that the quoted Turner lemmata are semantically underdifferentiated and seem to contain originally separate words. The Bangani and Khowar forms are especially close to the Armenian word which may be an indication that this is a loanword.

Rom.Arm. *santhu* 'oven'

Compare Bangani *sɔndaɔ* 'to warm or heat up (e.g., an aching limb through a hot compress)', and (poetic) *sɔdaɔ* and *sɔdar* 'funeral pile'. Turner postulates *sámdahati* 'burns up' (12899a) with one modern descendant in Maldivian. The Bangani forms, however, rather derive < OIA *sámdāhayati* 'to cause to burn up'. The Rom.Arm. word could be inherited, but the scarce evidence just from the fringes of Indo-Aryan makes borrowing not unlikely.

Rom.Arm. *samli karel* 'to fabricate, prepare, make, build'

OIA *sam̐māti* 'construct, grant' (12975) has the past participle *sámmīta-* 'of the same measure' (Pali *sammita-* 'measured'). The Armenian form must derive from the past participle. Regarding the meaning, cf. Phalūrā *samúm* 'I build, arrange (house, bed)' (sub 12975) and Bangani *sɔminuaɔ* 'to manufacture, make; to repair, join (together), fit together (as carpenter)'. The Bangani word is a synonym verbal compound with independent *nuaɔ* meaning 'to bend'. The Armenian form is certainly a loanword since the meaning 'construct' is limited to the north-west and the *-l-* seems to reflect an old *-ta-*.

Rom.Arm. *sol, söl* 'loud; voice, word, narration, news'; plus *lel* 'to take' it means 'to revile'; plus *ǵrel* 'to call, shout, sing' (which also has this meaning); *sol* plus *karel* 'to do' means 'to speak, narrate'

The forms are connected with OIA *ślóka-* 'sound, hymn of praise' (12748). Since modern descendants are limited to the north-west this must be a loanword.

Rom.Burg. *hišano* ‘shady’ and *hišo* ‘shade’, and perhaps Rom.Syr. *ausa* ‘shadow’

The forms are not connected with Romani *učhal* ‘shadow, shade’ which is < OIA **avacchāda-* ‘cover’ (763). Compare Indus Kohistani *išə̀n̄* ‘(a place which is usually) shady, without sun’. But further connections are (yet) unclear.

6 Possibly inherited but not yet (clearly) etymologized Romani words

Rom. *avertehara* ‘day after tomorrow’

Compare Waigali *varātr* ‘tomorrow’ which is < OIA *apararātra-* ‘latter part of the night, end of night’ (436). The Romani word was probably suffixed with an *-r*-extended form < OIA *áhar-* ‘day’ (993) as in Khowar *avēri/averā* ‘day after tomorrow’. The extension is also found in Marāṭhī *sāterē* ‘weather lasting for seven days’ which is < OIA *saptāhā-* ‘period of seven days’ (13161). The existence of *áhar-* is further corroborated by Rom.Dol. *prektaha* ‘übermorgen’ which consists of a Slavonic prefix and the inherited word.

Rom.Arm. *chasachutēn, chasautēn* ‘shame, shamefacedness’

The word contains a derivation < OIA **chupti-* ‘touch’ (5057) with modern descendants meaning ‘impurity’. Cf. Hindi *achūt* ‘(an) untouchable’ and *chuāchūt* ‘restrictions on touching, or contact’.

Rom.Arm. *nēnel* ‘to carry’

This is perhaps a derivation < OIA *nayana-*¹ ‘leading’ (6967) which is a nominalization of OIA *náyati* ‘carries off’ (6966). European Romani *anel* ‘to fetch, bring’ is < related OIA *ānayati* ‘leads forward, fetches’ (1174).

Rom.Arm. *panghri, pantry, panghyn* ‘hen, chicken’

The first form derives < OIA **pakṣirūpa-* ‘bird’ (7637), but the second and third forms are less clear as they seem to contain suffixes.

Rom. *merikli* ‘necklace, bracelet’

This is a Wanderwort, cf. Indus Kohistani *malγalāray* ‘a pearl necklace; a diamond (it is said that it emits light by itself at night)’, Pashto *marγalara* ‘a pearl’, Sogdian (Middle Iranian) *māryārt* ‘pearl’, Avar language (North Caucasian) *margal* ‘pearl’, Chaldaean *margal*, *maregale* ‘pearl’, Armenian *markarid* ‘pearl’, Persian *marvarid* ‘pearl’, Gothic *markreitas*, Walachian *margarita*, *merjeritarju*, Albanian *margaritar*, Greek *μαργαριτης μαργαρις*, *μαργελλιον*, *μαργηλις*, Latin *margarita* ‘lapis indicus’, etc.

Rom.Arm. *vahičq* ‘axe’

Regarding the Armenian suffix *-čq* see Finck (1907: 51); the lexeme is < OIA *vāši-* ‘sharp-pointed knife or adze’ (11588).

Rom. *šudro* ‘cool, cold’

The etymology proposed by Tálos, *šudro* < OIA **śuddha-ra-* ‘clear (water)’ < ‘cold (water)’ is not convincing (cf. OIA *śuddhā-* ‘clean, bright, white’ [12520]). Instead the word derives < OIA *túṣāra-* ‘cold’ (5894) with subsequent metatheses and voice assimilation of the original *-t-*. The metatheses must have occurred after leaving the north-west when the rule *-t- > -l-* was not active anymore.

7 On the original home of Romani

As has been pointed out by Turner (see section 3), Romani originally belongs to what he calls the Central group of Indo-Aryan. This is perhaps not exactly the same as Grierson’s division of New Indo-Aryan into three subgroupings which he called midland or inner languages, intermediate languages and outer languages, but it is also not very different. We have already seen that Romani has borrowed a substantial amount of words from the north-western branch, and below more data will be furnished. However, I want to stress in this section that Romani contains a substantial amount of words which are apparently limited to the inner languages. This might be taken as additional corroboration for the assumption regarding Romani belonging originally to the inner group. The following words appear to me as potential

candidates for the inner group as none of them has descendants in the outer languages:²⁰

Rom. *čunř/čunřa* ‘tress, plait’

One finds here some closely related OIA forms—*cūḍa*-¹, ‘protuberance on brick; topknot on head’, **conḍa*-, **coṭṭa*-, **cunda*- (4883)—with a main meaning ‘(knot of) hair’. According to Turner (who quotes Mayrhofer) this is a Dravidian lemma; but the problem is that its alloforms have quite different geographical distributions. This suggests overlap of more than one lemma:

cūḍa-¹: Nuristani, Dardic, West Pahārī, Sindhi, Lahnda, Panjabi, Kumaoni, Nepālī, Bengali, Assamese, Oriya, Bihari, Bhojpuri, Maithili, Hindi, Maldivian.

**conḍa* -: Dardic, Lahnda, Panjabi, Hindi.

**coṭṭa* -: Sindhi, Lahnda, Panjabi, West Pahārī, Hindi, Oriya, Gujarati.

**cunda* -: Oriya.

The relatively early attested *cūḍa*-¹ is also the most widespread allomorph with modern descendants found in the triangle between Nuristani, Assamese and Maldivian. The allomorph **conḍa*-, from which the Romani forms derive, is largely limited to the central languages (and Dardic).

Rom. *bilal* v.i. ‘to melt, thaw’ and *bilavel* v.t ‘to melt’

The words derive < OIA *vilīyate* ‘is dissolved, melts’ and *vilāpayati*² ‘dissolves, melts’ (11906). Modern descendants are predominantly found in Nuristani, Dardic, Central and Eastern

20 It is very likely that the inner (or central) languages exerted a heavy influence on the Pahārī languages and large parts of the north-western languages. Influence in the other direction was certainly much less pervasive. Thus, as long as there are no clear markers identifying a word as belonging to the north-west, the following examples having such parallels can be either borrowings from the central area into the north-west or they belong to a common heritage.

Pahari, and Central NIA.; but there are no descendants in the outer languages.²¹

Rom.Ger. *daro* ‘tree’

Not mentioned by Turner,²² < OIA *dāru-* ‘piece of wood’ (6298). Modern descendants are found in Dardic, Pahārī, Central NIA and Sinhalese. Thus this is a word shared by the inner and the north-western languages, but is absent in the outer languages.

8 Some close coincidences between Romani and Ḍumāki

Ḍumāki is an Indo-Aryan language spoken by the Ḍoma in the Hunza valley and some other scattered places in the Northern Areas of Pakistan. Ḍumāki is not a Dardic language—even though it has been heavily influenced by Dardic Shina, by an unknown other Dardic language and by Burushaski—but is related to languages of the North Indian plains. The language shows some peculiar coincidences with Romani which are worth to be quoted here:

Ḍumāki	Romani	Old Indo-Aryan
<i>karká</i> ‘bitter’	<i>kerko</i> ‘bitter’	<i>kaṭú-</i> , <i>kāṭuka-</i> ‘pungent, bitter’ (2641)
<i>phaaká</i> ‘shoulder’	<i>phiko</i> ‘shoulder(-blade)’	<i>*sphiyá-</i> , <i>sphyá-</i> *‘scapula’ (13839)
<i>g’i’ryu</i> ‘butter’	Ḍomari <i>gir</i> , <i>giri</i> ‘butter’	<i>ghṛtá-</i> ‘ghee’ (4501)
<i>pursúm</i> ‘flea’	<i>pušum</i> ‘flea’	<i>*pruṣi-</i> ‘flea’ + <i>maśáka-</i> ‘mosquito’ (9029 + 9917)

21 The lemma is also found in Sinhalese. But Sinhalese (including Maldivian) cannot simply be allocated to the outer languages. For the time being the question has to remain open.

22 See Siegmund Andreas Wolf *Grosses Wörterbuch der Zigeuner-sprache*.

Notes:

(1) In *phaaká* ‘shoulder’ the suffix *-ká* is remarkable as it is otherwise only found—besides Romani—in Kashmiri *phyok* and Iranian Shughni *fyak* and some other north-western languages (see above section 4).

(2) The European form for ‘ghee’ is *khil* and Shina has *gi*, but *Ḍumāki giríu* ‘ghee’ resembles the form in *Ḍomari*.²³

(3) The word *purṣúm* ‘flea’ is peculiar because of the final *-m* which is the rest of a second word, cf. Torwali “*p^ʷumāsh*” ‘flea’ (Barth and Morgenstierne) and Dameli *prāṣũ* ‘fly’. Apparently, either the first sibilant or the second was dissimilated due to syllable contraction.

9 Remarks on Norwegian and Swedish Romani

The first and the second lemma discussed in this section are related either through common etymology and/or through partly inextricable contaminations and borrowings, and (perhaps repeated) onomatopoetic creations. Yet they form a closely interrelated complex. Note, however, that the first lemma is limited to Scandinavian and German Romani whereas the second one is more widespread. The second lemma is said to derive from Greek and the first either from Turkish, from PIE or even Nostratic. However, both have Wanderwort characteristics. The second lemma has been included here because of its closeness to the first one.

kakni ‘hen’

Rom.N. *kakni* and Rom.S. *kàkkni* both ‘höna – hen’ (also *kákkno* = *kánno* ‘tupp – cock’)—but also Rom.Ger. (see Wolf 1960) *kachni*, *kaxni*, *kaghni* etc. all ‘Huhn, Henne’— are connected with Rom.

23 The similarity between the *Ḍumāki* and the *Ḍomari* words was already noted by Lorimer.

khajni ‘hen’.²⁴ Iversen (1944: 83) rightly reconstructs **kary-ni* with a parallel in Rom.Arm. *karyi* ‘cock’. The lemma is found mainly in Iranian and in north-west Indo-Aryan as well as dispersed in some other languages. Here follow some examples (for a more comprehensive list see Zoller forthcoming): Garhwālī *śyām karka* ‘woodcock’, Wakhi *kherk*, *khirk* ‘chicken’, Burushaski *qarqaámuc* ‘Huhn, Hahn’ (with *-muc* < OIA *mṛgacī* ‘bird’ [10265]),²⁵ Pashto *qārya* ‘crow, rook’, Ossetic *kark* ‘hen’, Late Avestan *kahrka-* (in compounds), Middle and New Persian *kark* ‘chicken, hen’, Tocharian B *krañko* ‘chicken’, etc. In addition the word is found in Modern Greek *karga* ‘a bird’ and Russian *kargá* ‘Krähe, Greisin’ etc. According to Doerfer (1967: 384) the original word is Turkish *qarya* ‘Krähe’. However, the reconstructed Proto Indo-European form is **kerk-*. But Mallory and Adams point also out (2006: 144) that onomatopoeic PIE **kVr-C-* probably not only provides the basis for ‘crow’ words but also for ‘hen’ words. Consequently the possibility cannot be ruled out that we deal here with a lemma that originates beyond the horizon of PIE, and in fact Nostratic origin is suggested by Nikita Krougly-Enke (p. 276–77, entry 6.20). He differentiates this from a similar lemma for magpie- or crow-like birds (p. 277, entry 6.21, see also next entry on ‘magpie’ etc.), but Rom.T. *karka* ‘magpie’ shows that there exist transitional forms between the two lemmata.

***kakkaráška* ‘big bird’**

Rom. *kakaráška* ‘magpie’, Rom.N. *kakkeraska* ‘magpie – skjære’, according to Iversen (1944: 82) Eilert Sund had noted the same form but in the sense of ‘eagle – ørn’, Rom.F. *kakaraška*, (*kakarachka*) and *kakkeraska* ‘Elster’, Rom.S. *kakkaráška*, *kàkkeraska* ‘skata, stor fågel, örn, falk, rovfågel – magpie, big bird, eagle, falcon, raptor’. There is Late Avestan *kahrk-āsa-* ‘vulture’ (see preceding entry); according to Cheung (2007: 168) this is a Wanderwort. He quotes among others Ossetic *cærgæs* ‘eagle’, Sogdian *carkas* ‘bird of prey’, Middle Persian *kargās*,

24 A probably older form is found in Rom.T. *kahnji* ‘hen’.

25 According to Gérard Fussman this is an onomatopoeic form (see 1972, entry 36).

Khwarezmian *krkys* ‘vulture’, New Persian *kargas*, Yidgha *karyəz*, Munjī *kargas* ‘black and white eagle’, Wakhi *karjpöps*, *karjöpc*, Sarīkolī *kargopc* both ‘magpie’, Sarīkolī *kiryī*, *kəγə* ‘small falcon’, *cory* ‘eagle’ as well as Old Indo-Aryan (late) *kr̥kaṣā-* ‘a kind of bird’. To this is to be added West Pahārī Rudhari (Kaul 2006 II: 270) *kəkras* ‘name of a grey-coloured ravenous bird with a long tail’. Here also belong Shina *kaṅkaróóco*, Yasin Burushaski *kuṅkuróco* and Burushaski *qamqurúúco* all ‘(Hahn) krähen’. According to Boretzky and Igla, Iversen and Tállos the word is of Greek origin. Iversen says (ibid.) “[t]he origin of this onomatope is to be found in Greek ‘karakaksa’ Elster (Pasp. P. 268); another, but less reasonable etymology, has been proposed by Bugge (p. 153). — The signif. given by Su. must be due to a mistake.” There is apparently also Turkish *kargas* ‘griffin’ and *kerkes* ‘phoenix’ (see e.g. Roelof van den Broek and Inez Wolf Seeger 1971: 204). The above evidence shows that the quoted authors are not wrong; however, *kakaráška* and similar forms are obviously the last ones of a series of repeated borrowings of a widespread and multiform Wanderwort that may, as suggested, have an origin beyond Proto Indo-European.

***grúmnin* ‘thunder’**

Rom.N. *gurmin* ‘thunder – torden’, Rom.S *grúmnin*, *gúrmin* ‘aska – thunder’, but also Rom.Ger. *grumos* ‘Donner’ and Rom.Dol. *grmini* ‘donnern’. There is Burushaski *qarám man-* ‘donnern’ which is also contained in Đumāki *lamqaram* ‘thunder’, and perhaps Shina *gram b-* ‘zusam-menstürzen’; Indus Kohistani *gharām-* *gharām* ‘sound of beating of the heart or of thunder’, Panjabi *gharamm* ‘splashing sound, sound of cannon or musketry’, Balochi *gran-* and Pahlavi **yarrān-* both ‘to thunder’. Even though there are lots of similar sounding words in NIA (usually of the form *gVr* as in Jaunsārī *garraṇō* ‘to growl’), the Romani word is usually seen as a borrowing from Slavonic **kurmí* or *grm(j)eti* ‘to thunder’. But the real background seems either to be an Indo-Aryan derivation of “North-West-PIE” **ghromos* ‘thunder’ or, what is more likely, there was repeated borrowing of similar sounding words.

***ghàna* ‘people’**

Rom.S. *ghàna* ‘folk, människor, bybor [icke resande] – people, folk, person, townsfolk (not travellers)’. Derivation < OIA *gaṇá-* ‘troop, flock’ (3988) is unlikely. A more likely derivation, already suggested by me (Zoller 2005: 151) is < OIA *ghaná-*² ‘compact, firm, dense’ (4424) with modern descendants in north-western languages like Bashkarīk *gān* ‘big, elder’, Savi *ghanyero* ‘elder’, Pašāī *gaṇ-ayōm* ‘my grandmother’. However, referring to Morgenstierne, Turner considers for some of the here quoted forms a reconstruction < OIA **ghaṇḍa-* as found in Iranian Parachi *ghaṇḍ* ‘big’ (in order to account for the retroflex nasal consonants). Here to be added are Indus Kohistani *ghō ṛh’ū* ‘elder/eldest brother’ (with second element < OIA *bhrātṛ-* ‘brother’ [9661]) and *ghō bā* ‘grandfather’ (with second element < OIA **bā-* ‘father’ [9198]). There is further connection between **ghaṇḍa-* and Kalasha *gáḍa* ‘big or mature of animate beings; elder’ which is used like Indus Kohistani in *gáḍa bába* ‘older sister’ and *gáḍa báya* ‘older brother’, but the loss of aspiration is inexplicable (Trail and Cooper’s connection with OIA *gāḍha-* ‘thick’ [4118] appears unlikely even though contamination cannot be excluded). There might be a further connection with Kalasha *gaṇḍāv* ‘statue of a deceased person’, and the basic semantics for the above words would thus roughly be ‘respected elder’, but again contamination by OIA *gaṇḍa-*² ‘trunk of tree’ (3998), and suggested by Trail and Cooper as the actual etymology, cannot be ruled out. Yet the Swedish Romani word is clearly a borrowing from a Dard language.

***čímpi* ‘turnip’**

Rom.S. *čímpi* ‘kålrot – turnip’ probably < OIA **chimba-* ‘pod, legume’ (12445) and probably inherited.

***čhòčha* ‘hare’**

Rom.S. *čhòčha* ‘hare’ is closely related with Rom. *šošoj* ‘hare, rabbit’ both of which are < OIA *śasá-* ‘hare’ (12357). The OIA lemma is found in all NIA branches; however note that allomorphs with an (sometimes aspirated) affricate instead of a sibilant and with *o* instead of *a* are limited to the north-west. Cf. Waigali *ścyuṇ*

and (Lumsden) “soče”,²⁶ Pašaī *čāska*, Jaunsārī *chaśā* all ‘hare’. Whereas the Waigali forms may reflect Indo-Iranian **časá-* (with metathesis), the Pašaī word displays an irregular development, and the Romani and Jaunsārī forms are Indo-Aryan. Still they seem to point to a common geographical origin in the north-west.

Matras’ suggestion that the *o* in the Romani word is due to masculine singular nominative ending in *-o* (2002: 39) is unsatisfactory. The affricatization in Rom.S. has parallels in Rom.S. *dóčhalo* ‘skyldig, ansvarig – guilty, responsible’ which belongs to Romani *doš* ‘fault, sin’ and which is < OIA *doša-* ‘fault’ (6587), Rom.S. *kass* or *kačh* ‘hö – hay’ which is < OIA *ghāsa-* ‘food, pasture grass’ (4471) etc. Moreover, there is also Rom. *čhela* ‘measles, (small)pox’ (discussed in section 4) with an affricatized *š-*, which is a borrowing from the north-west.

The case of Yenish (German: Jenisch)²⁷ *čukel* ‘sauer’ (besides *šukel* ‘sauer’, see Josef K. von Train) is not quite clear. A derivation < OIA *cukrá-* ‘sour, sharp to the taste’ (4850) is phonologically possible (compare Ashkun *čukalá* ‘sour, bitter’) but would seem to be without parallel in Romani. I therefore assume that also this is a case of affricatization and that the word derives < OIA **śuktala-* ‘become acid or sour’ (12504) like Romani *šuklo* ‘sour’. Thus it seems that some of these words underwent affricatization in the north-west and others, much later, perhaps under dialectal German influence.

***pall* ‘apple’**

There are Månsing²⁸ *pall* ‘apple – apple’ (Pa.), Nasare²⁹ *pall* ‘ditto’ and Knoparema³⁰ *pállar* ‘ditto’ (Pa.). Here probably also Yenish *balling* ‘Apfel’ (see Josef K. von Train). They are close to Nuristani Waigali *pal’ā* ‘Apfel’, Prasun *vā* and Kamviri *pař’ə*

26 The reference to Lumsden is found in Morgenstierne 1954: 154.

27 On this variety of German see http://de.wikipedia.org/wiki/Jenische_Sprache

28 According to Iversen (1944: 10) this is “... the secret idiom of the Westgothian pedlars (“knallarna”)...”

29 According to Iversen (ibid.) it is an “... idiom used by the “nasare”, a sort of pedlars corresponding to the German Hausierer...”

30 This is, according to Iversen (1944: 9), “... the idiom of the Swedish chimney sweepers.”

‘apple’, Dardic Gauro *paló* ‘apple’ and Šātōṭi *phalā* ‘an apple’ which is probably a borrowing from Shina *phalā* ‘ditto’; and Burushaski *báalt* ‘Apfel’. See Turner’s sceptical comments sub *phála*⁻¹ ‘fruit’ (9051) and *pāṭali*- ‘Bignonia or Stereospermum suaveolens’ (‘yellow snake tree’) (8034) regarding a possible derivation. There is no doubt that the above Scandinavian Romani ‘apple’ words are borrowings from the north-west. But there are also the following words which belong to the above-mentioned (section 4) Romani *phab*, *phabái* ‘apple’: Rom.N. *pabb*, rarely *babb* ‘potato’, Rom.F. *phab* ‘Apfel’, and Rom.S. *pabbar* ‘potatoes’ (which are also borrowings from the north-west).

I have no explanation for the fact that one Indo-Aryan word is found among the Scandinavian Roma speakers and the other among peddlers and chimney sweepers.

10 Fronting of aspiration (*h* metathesis)

In an article from 1959 (1975) Turner discusses fronting of aspiration in European Romani. He summarizes his conclusions at the end of the paper (1975: 388). The most important conditions for fronting are:

- When the initial consonant was *g, j, d, b* (*v*), *ś, p*.
- When the internal consonant or consonant group was *ggh*, *ṅgh*, *ṅkh* (> *ṅgh*), *jgh*, *cch* (?), *ḍḍh*, *ṭh* (> *ṛh*), *ddh*, *ndh*, *th* and *dh* (> *lh*), *bbh*, *mbh*, *ph* (> *bh, vh*).
- It did not occur when the internal consonants were *kkh*, *cch* (?), *kh* and *gh* (> *h*), *bh* (> *h*), *rh* (?).
- Apart from two exceptions, fronting of aspiration occurred only between a voiced initial stop and a voiced internal aspirate.

Moreover, Turner shows (1975: 381) that fronting of aspiration preceded the devoicing of voiced aspirates in European Romani. All the above points indicate that the process of fronting of aspiration in Romani occurred when its speakers stayed in the north-west of South-Asia. In other words, *h* fronting in Romani was caused by the same process as in Dardic (and perhaps West Pahārī). Here is further evidence from Romani not yet noted, however without known parallels in NIA:

khanči ‘nothing, something, anything’ < OIA *kimcid* ‘anything’ (3144) with *-*cch-* from OIA *kāścid*, cf. Bhalesi *kich* ‘some, a little’ (Turner mentions only Rom.Pers. *hič* ‘nothing’, Rom.G. *ič*, *hič* ‘something, nothing’, Rom.Wel. *čī* ‘anything’).

čhingar ‘row, quarrel, noise’ < OIA *ciṅghāṭa-* ‘noise, scream’ (4787) (Turner mentions only Rom.Wel. *čiṅār*, *čiṅarī* ‘row, quarrel, brawl’).

pherja(s) ‘fun, joke’ < OIA *parihāsa-* ‘jesting, ridiculing’ (7902) (Turner mentions only Rom. and Rom.Germ. *peryas*, Rom.Wel. *paia* ‘fun, sport, joke’).

Fronting parallels in the north-west

Since fronting preceded devoicing, it is not surprising to find in the north-west several parallels of *h* fronting but (so far) no parallel for *h* fronting *plus* devoicing. This may be an indication that devoicing in Romani occurred outside the South Asian language area, namely in Armenia³¹ (see e.g. Donald Kenrick 2004: 31). Note, however, that the proto-forms for the following modern Dardic and the Romani words are sometimes not identical but only closely related:

31 Of course this did not include Ḍomari.

Meaning	Romani	Dardic	MIA	OIA
armpit	<i>khak</i>	<i>khačāl</i> ³²	<i>kakkha-</i> <i>kaccha-</i>	<i>kákṣa-</i>
smell	<i>khand</i>	<i>ghond</i> ³³	<i>gandha-</i>	<i>gandhá-</i>
pregnant	<i>khabnī</i>	<i>ghabī</i> ³⁴	<i>gabbhinī-</i>	<i>garbhīnī-</i>
to weave	<i>khuvel</i>	<i>ghum</i> ³⁵	<i>gu(b)haī</i>	<i>guphāti</i> <i>gumphāti</i>
tongue	<i>čhib</i>	<i>žhip</i> ³⁶	<i>jibbhā-</i>	<i>jihvā-</i>
month	<i>čhon</i>	<i>jhūn</i> ³⁷	<i>joṇhā-</i>	<i>jyótsnā-</i>
moon	<i>čhomut</i>			
and	<i>thaj</i>	<i>the</i> ³⁸	<i>*tahavi</i>	<i>tathāpi</i>
(molar), beard	<i>thar</i>	<i>d^hāri</i> ³⁹ <i>dhāi</i> ⁴⁰ <i>dharī</i> ⁴¹	<i>dāṭhā-</i>	<i>dāmṣṭra-</i>
wing	<i>phak</i>	<i>phāṣⁱēl</i> ⁴² <i>phačāl</i> ⁴³ (<i>< *pačh</i>) ' <i>phači</i> ⁴⁴ <i>phakhru</i> ⁴⁵	<i>pakkha-</i>	<i>pakṣá-</i>
to bind	<i>phandel</i>	<i>bhónik</i> ⁴⁶	<i>bandhati</i>	<i>bandhati</i>

32 Bhatise. Turner mentions only Rom.Wel. *kakh* and Rom.G. *kak*.

33 Kalasha.

34 Indus Kohistani.

35 Kalasha with meaning 'multi-stranded bunch (of beads)'.

36 Phalūrā and Kalasha.

37 Jaunsārī. Regarding the ending in *čhomut* cf. Satlaj Group *jōth* 'moon' which is discussed above in section 5.

38 Khowar. See comments in section 4.

39 Pašāi.

40 Maiyā. See Fussman 1972, entry 19, but aspiration is not confirmed in Zoller 2005.

41 Parachi.

42 Bhatise.

43 Shina.

44 Dameli and meaning 'bird', thus < related OIA *pakṣīn-* 'bird' (7636).

45 West Pahārī Rāmbani with meaning 'birds' and thus going back to OIA **pakṣirūpa-* 'bird' (7637).

46 Kalasha.

to ask	<i>phučel</i>	<i>phučik</i> ⁴⁷	<i>pucchati</i>	<i>pr̥cchāti</i>
--------	---------------	-----------------------------	-----------------	------------------

Note: Romani *khul* ‘excrement’ is < OIA *gūtha-* ‘excrement’ (4225) and there are a few NIA cases with fronting as in Lahnda *ghū* and Oṛiyā *ghua* both ‘excrement’. But these appear to be independent developments.

The phenomenon of *h* fronting is much more comprehensive in Dardic and West Pahārī than it is in European Romani and it allows fronting of basically all aspirated sounds, single ones and clusters. On the other hand, in Romani it rather appears like an incomplete phonological process. This can only mean that Proto-Romani indeed was affected by *h* fronting during the period when its speakers stayed in the north-west of South Asia, but apparently the process was in the first millennium CE not as advanced as we find it today. Recalling Turner’s observation that apart from two exceptions, fronting of aspiration occurred only between a *voiced* initial stop and a *voiced* internal aspirate in Romani it is now clear that this exactly reflects the situation in Dardic and West Pahārī where one observes a loose “connection” between *h* and its associated voiced stops as against the case of aspirated unvoiced stops (see Zoller forthcoming) as e.g. in Indus Kohistani nominative *ghūi* ‘mare’ but ergative *guhê* (< OIA *ghoṭa-* ‘horse’). Turner therefore rightly concludes (1927: 304) that *h* fronting in case of aspirated tenuis occurred later than that of aspirated mediae. Exactly this has also to be presumed for Dardic and West Pahārī.

11 Aspiration not justified by etymology

This phenomenon, which was first reported by Morgenstierne for Phalūrā, is in fact widespread both in Dardic and in West Pahārī (see Zoller forthcoming), e.g. Dardic Indus Kohistani *thók^h* ‘clod’ < OIA **tukka-* ‘piece’ (5466) or West Pahārī Jaunsārī *bhirai* ‘cat’ < OIA *bīḍāla-* ‘cat’ (9237). However, it is also known from

47 Kalasha.

Middle Iranian Sakian as in *phattanai* ‘palate’ < IIR *paθana-* ‘broad, wide’, *phara-* ‘much’ < IIR *paru-*. There are also a few Indo-Aryan instances in European Romani, which testify that they are borrowings either from Dardic or West Pahārī. All this is not unlikely as the phenomenon apparently occurred already occasionally in MIA Gāndhārī as in *dhakṣiṇami* for OIA *dākṣiṇa-* ‘southern’ (6119) (see Fussman 1989: 482 and Salomon 2002: 132):

čhamb ‘skin (on fruit), rind (bacon)’ < OIA *cārman-* ‘hide, skin’ (4701),

phumb ‘pus’ < OIA *pūya-* ‘pus’ (8328),

phurano ‘old’ < OIA *purāṇá-* ‘ancient’.

Note that Romani has borrowed at least one Iranian word displaying aspiration not justified by etymology: *phurd*, *phurt* ‘bridge’ ← dialectal Iranian *phurd* ‘bridge’.

12 *r* metathesis (fronting)

There are not that many cases for *r* metathesis in Romani; still, also this process is probably due to contact with languages in the north-west. Besides Dardic and West Pahārī, and also Sindhi and Lahnda, it is again also found in Middle Iranian Sakian (but apparently not in Nuristani), as the following examples demonstrate: *grāma-* ‘hot’ < IIR *garma-*, *drays-* ‘hold s.o./s.th.’ < IIR *darz-*, *drāmma-* ‘pomegranate’ < IIR **darma-* (OIA *dāḍima-*), *drūba-* ‘name of a certain plant’ (perhaps an IA borrowing, cf. OIA *dūrvā-* ‘the grass *Panicum dactylon*’ [6501]), *brumja-* ‘birch bark’ (again perhaps an Indo-Aryan borrowing, cf. OIA *bhūrja-* ‘birch tree’ [9570]). An example from Indus Kohistani is *ṛūb* ‘the grass *Panicum dactylon* or a similar variety’ (with *ṛ-* < older *dr-*, cf. Sakian *drūba-*) < OIA *dūrvā-* ‘the grass *Panicum dactylon*’ (6501) and an example from West Pahārī is Khashi *bhrebhu* ‘brown bear’ < OIA *babhrú-* ‘reddish brown’ (9149). This *r* fronting was already known in Gāndhārī as seen in *dhrama-* instead of OIA *dhárma-* (6753) (see Fussman 1989: 487). As *r* fronting is found over a considerable geographical area in the west and north-west,

it is difficult to assign a specific place from where the following Romani words have been borrowed:

breš ‘rain’ < OIA *varṣá-* ‘rain’ (11392)

tradel ‘to drive (away)’ < OIA *tárdati*¹ ‘sets free’ (5721)

pravarel ‘to nourish s.o.’ ← Persian *parvardan*

However, a parallel is found between Romani *prašav* ‘rib’ (actually ‘part of a cauldron’,⁴⁸ regarding ending cf. Rom.Wel. *pāšavo* ‘rib’) and Khowar *praš*, Savi *prašú*, Phalūrā *prāšū*, Shina of Gilgit *prāšǰ* all ‘rib’, and probably also Kalasha *praš* ‘steep hillside’, all of which are < OIA *pársu-*¹ ‘rib’ (7948).

13 The historical change *a* > *e*

Yaron Matras writes (2003: 34): “Preceding simple consonants, historical /a/ is represented in Romani by /e/, and in some cases by /i/: OIA *kar-*, Romani *ker-* ‘to do’... Historical /a/ is retained however in positions preceding an historical consonant cluster: OIA *gharma*, Romani *kham* ‘sun’...” Somehow contradicting this statement he quotes on p. 39 OIA *varṣa* > Romani *berš* ‘rain’. Similar vowel changes *a* > *e* are also known from the north-west, however, there they are only occasionally found and in a scattered way. Moreover, it seems that the vowel changes were caused by different factors, e.g. epenthesis or shift of accent. Thus there are quite many cases of epenthesis in Nuristani, Dardic and West Pahārī where a final *i* causes an initial *a* to be raised to *e* as e.g. in Ashkun *vērī* ‘speech, language, word’, Kati *verí*, Prasun *verī* which are < OIA **vari* ‘speech’ (11327), Bhatise *čhⁱēl* ‘a goat’ < OIA *chagalikā-* ‘goat’ (4963), Bashkarīk *ǰēng* ‘shin-bone’ < OIA **jaṅghiya* ‘belonging to the shank’ (5084), etc. I will therefore not try to formulate alternative rules but simply list the cases where I see possible parallels between Romani and north-western languages. Hence, these parallels suggest that the Romani vowel

48 See <http://ling.kfunigraz.ac.at/~rombase/cgi-bin/art.cgi?src=data/ethn/work/copper.en.xml>

change *a* > *e* occurred through contact with north-western languages, as it would otherwise fail to explain why this rule did not apply in non-IA borrowings.

ivend ‘winter’ is usually derived < OIA *hemantá* ‘winter’ (14164); however, compare Pašai (some dialects) *emén*, *emen* ‘winter’ which Turner derives < OIA *hemānta*. But there is no phonological motivation for the *e* in the second syllable and thus it is not clear whether we have here an accidental similarity.

kerel ‘to do, make’ < OIA *karóti* ‘does’ (2814) – compare the conjugation of Indus Kohistani *karáv* ‘to do’ which displays “irregular” variations of the vowel e.g. in present continuous *su^h kèrã bét^h* ‘he is doing (s.th.)’; Bhadrawāhī *kernū* ‘to do’, Kuḷuī *kēnu* ‘to do, make’. The *e* may be due to epenthesis, compare Niya documents *kareti*, Prakrit *karēi*, *karai* ‘does’ and Tirahi present tense *karēm*, °*ēs*, °*ē*.

gelo ‘went, gone’ is preterite of *džal* ‘to go’ and derives < OIA *gatá* ‘gone’ (4008) – compare the Rambani dialect of Kashmiri *gēu*, Panjabi *geā*, Bhalesi *geū* and Khashi *gedo* ‘went’.

dženo ‘(male) person’ < OIA *jána* ‘race, person’ (5098) has an exact parallel in the Urtsun variety of Kalasha *žen* ‘person’.

deš ‘ten’ < OIA *dáśa* (6227) – compare Indus Kohistani *dlyšⁱ*, Prasun *lez*, Shina *dái* all ‘ten’ which point to a Dardic and Nuristani proto-form **daši*, although the final vowel is inexplicable.

berš ‘year’ < OIA *varšá-* ‘year; rain’ (11392) – compare Phalūrā *beriš*, Shina of Gilgit *bəriš* and Bangani *bərīš* all ‘year’ which are < older **bariśa* (see Prakrit *varisa* ‘rain’ and Kashmiri *varih* ‘year’).

men ‘neck’ < OIA *maṇi* ‘hump of camel’ (9732) has no comparable modern parallel in South Asia; however, the vowel is probably the outcome of epenthesis.

mel ‘dirt’ is usually derived < OIA *māla* ‘dirt’ (9899) but a derivation < OIA **malin* ‘dirty’ (9904) is equally possible and would explain the Romani vowel; cf. under 9904 Khashali *mel* ‘ear-wax’, Kotgarhī *meḷa* ‘dirt’, Garhwālī *mel* ‘dirt’.

len ‘river’ < OIA *nadī* ‘river’ (6943) with consonant metathesis – compare Woṭapūrī *nyed*, Gawar-Bati *nēndi*, Torwali *ned*, Ashkun *nēd’i*, *n’ēdī*, Dameli *nalī*, Savi *neli*, Chameali *nei* all ‘stream’. Thus there is also here a case of epenthesis. Note that Rom.T. *ljen* ‘river’ seems to have preserved the word-final palatal vowel.

les ‘him’ < OIA *tá* base of nominative singular neuter, genitive masculine, neuter *tásya* (5612) – compare Kalasha *te* ‘they’; Torwali *tes*, Bhalesi *tes*, Bangani *tes*. Here we have again a clear case of north-western epenthesis. But the change of initial *t-* > *l-* in the Romani form is unusual and has only one parallel in *len* ‘river’. But there one would expect the change *-d-* > *-l-* to have occurred before the metathesis even though there is no supporting evidence from Dardic. Could the word have been borrowed from Rom.Arm.?

šel ‘hundred’ < OIA *śatá* ‘100’ (12278) – compare Bhatise *ś’āl*, but there are also other NIA forms with an *i* as in Nepālī, Bihari, Marathi *sai*, Sinhalese *siya*, etc.; compare above *deś* ‘ten’.⁴⁹

There are also some words where the *e* goes back to older *i*. Since the number of examples is so small it is, however, an open question whether or not this sound change in Romani was caused by an influence from north-western languages:

šero ‘head’ < OIA *śiras* ‘head’ (12452) – cf. Kashmiri *hēri* ‘above, upstream’, Siraji of Doda *šeri* ‘head’ (Kaul 2006 II: 328).

49 On *šelī* ‘bran’ and Rom.G. *seli* ‘bran’ see above section 4 where I have suggested an effect of epenthesis.

čhela ‘smallpox’ (discussed above in section 4): compare Waigali *šéle* and Woṭapūrī *šēl* ‘cold’.

14 Compound building

There is the Romani adjective *šernango* ‘bare-headed’ which is a compound consisting of *šero* ‘head’ and *nango* ‘naked’. There are also the following compounds with first element a body part followed by a qualifying adjective: *kašuko* ‘deaf’ probably built with *kan* ‘ear’ and *šuko* ‘dry’, and *punřango* ‘barefoot’ which is built with *punřo* ‘foot, leg’, a side-form of *pinřo* ‘ditto’, plus again *nango*. These words look like loan formations after a Dardic model, compare Indus Kohistani *šiš-lúṭṭ*^h and Phalūrā *šiša-lúto* both ‘bare-headed’ (second element < OIA **luṭṭa/luṭṭha* ‘defective’ [11076]; also in Shina *lúto*, Burushaski *lóto* ‘barhäuptig, entblößt, schamlos’). Nominal head-modifier constructions with the adjective following the noun are quite common for instance in Indus Kohistani.

15 The -do Suffix

We have already come across this suffix above in section 4 in Romani *bašaldo* ‘musician’, Rom.T. *bašado* ‘violinist’ and in Indus Kohistani *gildò* ‘singer’, and I have pointed out that Rom.S. *ghildo* ‘fest, party’ must formerly have had the same meaning. The second element *-do* derives < OIA *dádḥāti* ‘places, lays on, gives, seizes’ (6145) but with the meaning ‘produce/belt out/perform a song’. The ending has nothing to do with the homophonous Romani participle as claimed by Boretzky and Igla (1994: 411)—compare e.g. *mardo* ‘beaten’—with the agent nouns here under discussion. There are at least two more Romani words with the same suffix: *lurdo* ‘soldier’ is an agent noun connected with *lurel* ‘to rob’ which is < OIA **luṭṭati* ‘plunders’ (11078), and *kheldo* ‘player’ which is also an agent noun connected with *khelel* ‘to play, dance’ which derives < OIA **kḥel-/kḥell-* ‘play’ (3918).

16 The conjunct verb *del*

In Romani one finds a number of verbal expressions consisting of a noun and the verb *del*. It seems so that in some of these expressions *del* goes back to OIA *dádāhāti* ‘places, lays on, gives, seizes’ (6145) while in other cases it goes back to the very similar OIA *dádāti* ‘gives’ (6141). In any case it is not always clear whether a conjunct verb has been inherited or borrowed in the north-west. Yet, a north-western influence seems highly probable in case of several of the following expressions:

kandel ‘to obey’ is probably a conjunct verb built with *kan* ‘ear’, cf. Hindi *kān denā* ‘to listen’ and Kalasha *kō kárik* ‘to listen to, to obey’ (with *kárik* < OIA *karóti* ‘does’ [2814]).

xanřudel ‘to scratch, scrape’ is probably a conjunct verb with first element < OIA *kaṇḍū-* ‘itching, the itch’ (2688) and has a parallel in Indus Kohistani *kán diyáṽ* ‘to scratch (o.s.), itch’.

del čhik or *čikdel* ‘to sneeze’ has usually NIA parallels with the verb directly derived from the noun *chikkā-* ‘sneeze’ (5032) as in Hindi *chīknā*. However, conjunct formations are again found in Dardic as in Gauri *čhígi ghō* ‘to sneeze’ with the verb deriving < OIA *ghátate* ‘is busy with’ (4407).

čučidel ‘to breastfeed’ with the noun deriving < OIA **cuccu-* ‘female breast, nipple’ (4855) has a direct parallel in Kalasha *čúču dek* ‘to breastfeed’ (cf. Hindi *dūdh pilānā*).

čumi-del ‘to kiss’ has conjunct verb parallels e.g. in Hindi *cumban karnā* ‘to kiss’ and in Bangani *khubi dennō* ‘to kiss’ (lit. ‘to give a kiss’).

boldel ‘to turn, return, bring back’ is a conjunct verb as Tálos’ reconstruction **bolə²* ‘rotates tr’ shows. I agree with him that *bol-* derives < OIA *valati, valate* ‘turns, turns to, speeds towards’ (11405), but his claim that the Romani form actually derives < the

weak form *ulyate* is phonologically difficult and would be without parallel. There is another parallel in Indus Kohistani *bál karáv* ‘to swing (on a swing)’ (with *karáv* < OIA *karóti* ‘does’ [2814]) and in Burushaski *bíli man-* ‘schaukeln’, *balbalán* ‘herunterrollen’ etc.

brišind del ‘it rains’ has a parallel in Indus Kohistani *λz diyáv* ‘to rain’ (but with first element < OIA *abhrá-* ‘rain’) and in Kalasha *bášik dyek* ‘to rain’ – cf. the differently formed Hindi *bāriś hotī hai, barsnā*.

late-del (Sinti) ‘to kick’ is a conjunct formation with first component deriving < OIA **lattā* - ‘foot, kick’ (10931); it has a parallel in Indus Kohistani *lāytav diyáv* ‘to kick’ and probably also in Bangani *lotti dennɔ* ‘to clear off, push off’. Less close is Hindi *latti mārṇā* ‘to kick’ or even *latiyānā* ‘ditto’.

šol del ‘to whistle’ (discussed above in section 4) has exact parallels in Indus Kohistani *šur-šurī diyáv*, Shina *šurúki d-* and Bangani *sēr dennɔ* all ‘to whistle’, but also in Hindi *sītī denā* ‘to whistle’.

jiv del (Rom.T.) ‘it is snowing’ (with first word < OIA *himá-* ‘cold, frost, snow’ [14096]) can be compared with Kalasha *kírik dyek* ‘to snow’ (with first word < OIA *kiri-* ‘scattering, heap’ [3175]) and also with Indus Kohistani *húū riṅgáv* ‘to snow’ (lit. ‘snow be attached’) but hardly with Hindi *barf parṇā* ‘to snow’ (lit. ‘snow to fall’).

17 Mythology

Magpie

Regarding Romani *kakaráška* ‘magpie’, discussed above in section 9, Iversen writes: “Mention may be made of the fact that for the older “travellers” this bird played a special part, as it was a prophetic bird from which they took auguries. A similar superstition is also found among English Gipsies...” However, this

must not be something specific to Roma traditions since the magpie has been regarded in Germanic mythology as a messenger of the gods or of the goddess of death *Hel* (see Wolfgang Epple). Moreover, since the word designates different birds we cannot know whether the Roma word previously also referred to the magpie. It is therefore difficult to say when and where the Roma adopted this mythologem.

Porcupine

According to Hermann Berger (1985: 793), there is a being in Roma mythology called *Hagrin* which is a „Dämonische[s] Wesen in Gestalt eines gelblichen Stachelschweins von ½ m Länge und 1 Spanne Breite. H. quält Tiere im Schlaf, besonders solche, die gerade geworfen haben, indem er sich ihnen auf den Rücken setzt und seinen Urin daran herabfließen lässt, wodurch eiternde Geschwüre entstehen“ (my rendering: The *Hagrin* is a demonic being in the shape of a porcupine with a length of ½ m and a breadth of 1 span. H. afflicts sleeping animals, especially those which have just delivered, by sitting on their back and by having its urine running down over them, through which purulent abscesses develop). A parallel is found in Shina and Burushaski *hargín* ‘Drache, Ungeheuer, entsteht aus einer gewöhnlichen Schlange, wenn sie groß und alt wird’ (‘dragon, monster, evolves from an ordinary snake when it becomes big and old’). The description of the *Hagrin* also resembles strikingly what Karl Jettmar (1975: 285) says about a constrictor in the Hindukush: „Ferner hören wir von einer Riesenschlange mit goldener Mähne, *Hargin* genannt ...“ (‘In addition do we hear of a constrictor with golden mane and called *Hargin*’). Unfortunately Jettmar does not tell us more about this strange creature, but not only do both beings have basically the same name, but the porcupine is of yellowish colour which corresponds with the golden mane of the constrictor. It is thus likely that the ancestors of the Roma acquainted themselves with this mythological being in the north-west of South Asia.

Conclusions

The data presented in this article provide ample support for some important theses of Turner which have been quoted above in section 1: Romani belongs originally to the inner branch of Indo-Aryan; the speakers of early Romani left their original home (probably located somewhere in the Ganges valley) before the time of the Ashoka inscriptions; after leaving their original home the speakers of early Romani came in contact with speakers of Indo-Aryan languages in the north-west of the subcontinent, that is with speakers of Dardic, Nuristani and, perhaps, West Pahārī.⁵⁰ This contact must have continued over several hundred years before the Roma left South Asia.⁵¹ There have been different views whether or not the languages of the north-west left clear-cut traces in Romani. As pointed out above, Turner was sure that some north-western words were borrowed into early Romani. Matras, however, is sceptical and finds that the "... lexical evidence remains marginal and largely inconclusive. Noteworthy is the fact that there are hardly any phonological innovations that are shared with the North-western languages..." (2002: 47). However, elsewhere⁵² Matras admits north-western influence, and as an example he refers to the Romani pronominal suffixes (as in the Romani past-tense forms *kerdjo-m* 'I did', *kerdja-s* 'he/she did' etc.). The geographical extent of languages with pronominal suffixes, however, does not coincide with my postulated north-western branch. The north-western branch comprises Nuristani, Dardic and West Pahārī whereas pronominal suffixes are found in Kashmiri, Shina, Lahnda, Sindhi, Pögulī and some dialects of Panjabi. In

50 It is self-evident to presume that in the second half of the first millennium the differences between the Dard languages and the varieties of West Pahārī were less developed than they are now.

51 It goes without saying that these linguistic matters say nothing about the ethnic, social, cultural etc. features of the speakers. It is thus also impossible to say whether the speakers of Romani at the time of their leaving the subcontinent were the progeny of those who left their original home. Such questions are anyway not part of this article.

52 http://romani.humanities.manchester.ac.uk/downloads/2/Matras-Rmni_ELL.pdf (second page) and Matras 2009: 11–3ff.

other words, it is likely that early Romani adapted this pattern in the north-west, but it is not a pattern characteristic of the north-western branch of Indo-Aryan. Note also that although early Romani has been in long contact with languages in the north-west there are only few traces of depalatalization—that is phonetically e.g. the change of [tʃ] to [ts]—even though depalatalization is widespread and old in that area. However, there are some Nuristan, Dard and West Pahārī languages where depalatalization is not found. Thus we may assume that in the first millennium AD this process was less comprehensive than it is today. There is no doubt that the strongest influence on early Romani came through contact with languages of the north-western branch, and the influence is found on all levels of grammar. In addition, there are even a few traces of mythologems originally located in the north-west.

Abbreviations

<	historically deriving from	NIA	New Indo-Aryan
>	historically developing into	OIA	Old Indo-Aryan
←	borrowed from	PIE	Proto Indo-European
Ir	Indo-Iranian		
MIA	Middle Indo-Aryan		

Variants of Romani

Note that the following list of abbreviations is only meant as an approximation; mostly it does not refer to variants known under specific names.

Rom.	Romani	Rom.Germ.	German
Rom.Arm.	Armenian Romani		Romani
Rom.As.	Asian Romani	Rom.H.	Hungarian
Rom.Burg	Burgenland Romani		Romani
Rom.Dol.	Dolenskji Romani	Rom.Lov.	Lovari
Rom.Eur.	European Romani		Romani
Rom.F.	Finnish Romani	Rom.N.	Norwegian
Rom.G.	Greek Romani		Romani

Rom.Pers.	Persian	Rom.T.	Transylvanian
	Romani		Romani
Rom.S.	Swedish	Rom.Wel.	Welsh
	Romani		Romani
Rom.Sp.	Spanish	Rom.Zak.	Zakopane
	Romani		Romani
Rom.Syr.	Syrian		
	Romani		

Lesser known languages and their affiliations

Ashkun	Nuristani	Kalasha	Dardic
Bangani	West Pahārī	Kanashi	Western Himalayish
Bashgalī	Nuristani		(Tibeto-Burman)
Bashkarīk	Dardic	Kannauri	Western Himalayish
Bauri	West Pahārī		(Tibeto-Burman)
Bhadrawāhī	West Pahārī	Kashmiri	transitional between
Bhalesi	West Pahārī		West Pahārī and
Brokskad	Dardic		Dardic
Chilīs	Dardic	Kati	Nuristani
Chitkuli	Western Himalayish	Khaśdhari	a variety of Bangani
	(Tibeto-Burman)	Khashi	West Pahārī
Dameli	mixed Nuristani and	Khashali	West Pahārī
	Dardic	Khetrānī	Outer Indo-Aryan
Deogari	West Pahārī	Khowar	Dardic
Ḍogrī	transitional between	Kocī	West Pahārī
	W. Pahārī and Panjabi	Koṭgarhī	West Pahārī
Ḍumāki	Central Indo-Aryan	Kuḷuī	West Pahārī
Gambīrī	Nuristani	Maiyā	old designation for
Gari	Western Himalayish		Indus Kohistani
	(Tibeto-Burman)	Niṅgalāmī	Dardic
Gauro	a variety of Indus	Paṅgwālī	West Pahārī
	Kohistani	Parachi	Iranian
Gawar-Bati	Dardic	Paśāī	Dardic
Indus Kohistani	Dardic	Phalūṛā	Dardic
Inner Sirājī	West Pahārī	Pōgulī	West Pahārī
Ishkashmi	Iranian	Prasun	Nuristani
Jaunsārī	West Pahārī	Roshani	Iranian

Rudhari	West Pahārī	Shumashti	Dardic
Sainjī	West Pahārī	Torwali	Dardic
Sarīkolī	Iranian	Tregami	Nuristani
Satlaj Group	West Pahārī	Waigalī	Dardic
Savi	Dardic	Wakhi	Iranian
Šāṭōṭī	a variety of Indus	Woṭapūrī	Dardic
	Kohistani	Yazghulami	Iranian
Shina	Dardic	Yidgha-Munjī	Iranian
Shughni	Iranian		

Literature

- Aichele, W. 1957. Die ursprünglich aspirierten Medien der Zigeunersprache. In: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 10: 101–107.
- Ambrosch, Gerd and Dieter W. Halwachs. 2002. *Wörterbuch des Burgenland-Romani (Roman). Roman – Deutsch – Englisch; Deutsch – Roman; Englisch – Roman*. Unter Mitarbeit von Katharina Deman, Ursula Glaeser, Michael Wogg. Arbeitsbericht 10 des Romani-Projekts. Verein Roma/ Oberwart: <http://romani.uni-graz.at/romani/>
- Baart, Joan L. G. 1997. *The Sounds and Tones of Kalam Kohistani: with Wordlist and Texts*. Islamabad: National Institute of Pakistan Studies and Summer Institute of Linguistics. (Studies in Languages of Northern Pakistan 1).
- Bailey, Harold Walter. 1979. *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.
- Bailey, T. Grahame. 1908. *The Languages of the Northern Himalayas Being Studies in the Grammar of Twenty-six Himalayan Dialects*. London: Royal Asiatic Society.
- 1920. *Linguistic Studies from the Himalayas Being Studies in the Grammar of Fifteen Himalayan Dialects*. London: Asiatic Society Monographs, Vol XVII.
- Bailey, Thomas Grahame. 1924. *Grammar of the Shina (ṣiṇā) Language, Consisting of a Full Grammar, with Texts and Vocabularies of the Main or Gilgiti Dialect and Briefer Grammars (with Vocabularies and Texts) of the Kohistani, Guresi and Drasi dialects*. London: The Royal Asiatic Society. (Royal Asiatic Society Prize Publication Fund 8).
- Bakker, Peter and Yaron Matras. 2003. *Bibliography of Modern Romani Linguistics*. Amsterdam: John Benjamins.

- Barth, Fredrik and Georg Morgenstierne. 1958. Vocabularies and specimens of some S. E. Dardic dialects. Offprint from *Norsk Tidsskrift for Sprogvidenskap* 18: 118–136.
- Berger, Hermann. 1985. Mythologie der Zigeuner. In: Walter Haussig (Hrg.), *Wörterbuch der Mythologie. Bd. 5: Götter und Mythen des indischen Subkontinents*. Stuttgart: Klett-Cotta, pp. 775–826.
- 1998. *Die Burushaski-Sprache von Hunza und Nager; Teil III: Wörterbuch Burushaski-Deutsch, Deutsch-Burushaski*. (Neuindische Studien 13). Wiesbaden: Harrassowitz.
- Bielmeier, Roland. 1977. *Historische Untersuchungen zum Erb- und Lehnwortschatzanteil im ossetischen Grundwortschatz*. Frankfurt a. M.: Lang.
- Boretzky, Norbert. 1992. Zum Erbwortschatz des Romani. In: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 45: 227–251.
- 2005. Metathesis and other, functionally related, sound changes in Romani. In: Schrammel, Barbara, Dieter W. Halwachs and Gerd Ambrosch (eds.), *General and Applied Romani Linguistics: Proceedings from the 6th International Conference on Romani Linguistics*, pp. 128–143. München: Lincom.
- Boretzky, Norbert and B. Igla. 1994. *Wörterbuch Romani – Deutsch – Englisch für den südosteuropäischen Raum*. Wiesbaden: Harrassowitz.
- Bronkhorst, Johannes and Madhav M. Deshpande (eds.). 1999. *Aryan and Non-Aryan in South Asia: Evidence, Interpretation and Ideology*. Harvard Oriental Series, Opera Minora Vol. 3. Cambridge (Mass.): Harvard University.

- Buddruss, Georg. 1960. *Die Sprache von Woṭapūr und Kaṭārqaḷā: Linguistische Studien im afghanischen Hindukusch*. Bonner Orientalistische Studien, N. S. 9. Bonn: Selbstverlag des Orientalischen Seminars der Universität Bonn.
- 1967. *Die Sprache von Sau in Ostafghanistan: Beiträge zur Kenntnis des dardischen Phalura*. Münchener Studien zur Sprachwissenschaft, Beiheft M.
- 1975. Gāndhārī-Prākṛit *chada* „Ton“. In: *Studien zur Indologie und Iranistik* 1: 37–48.
- Cech, Petra und Mozes Heinschink. 2002a. Vokabular der Dolenjski Roma aus Novo Mesto und Bela Krajina, Slowenien. *Grazer Linguistische Studien* 58 (Herbst 2002) 1: http://www.uni-graz.at/ling2www_gls58_cech.pdf
- 2002b. Wörterbuch der Romani-Variante der Österreichischen Lovara. Lovari – Deutsch – Englisch. *Arbeitsbericht 3A des Romani-Projekts*, hrsgg. von Dieter W. Halwachs, Verein Romano Centro / Wien 2002: <http://romani.uni-graz.at/romani/publ/lovara.de.shtml>
- Cheung, Johnny. 2007. *Etymological Dictionary of the Iranian Verb*. Leiden, Boston: Brill.
- Degener, Almuth. 1998. *Die Sprache von Nisheygram im afghanischen Hindukusch*. Wiesbaden: Harrassowitz.
- Doerfer, Gerhard. 1967. *Türkische und mongolische Elemente im Neupersischen: Unter besonderer Berücksichtigung älterer neupersischer Geschichtsquellen, vor allem der Mongolen- und Timuridenzeit*. Wiesbaden: Franz Steiner.
- Elfenbein, J., D. N. MacKenzie and Nicholas Sims-Williams (compiled and edited). 2003. *A New Etymological Vocabulary of Pashto (by Georg Morgenstierne)*. Wiesbaden: Dr. Ludwig Reichert.

- Emmerick, Ronald E. and Prods O. Skjærvø. 1987. *Studies in the Vocabulary of Khotanese II*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- 1997. *Studies in the Vocabulary of Khotanese III*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Epple, Wolfgang. 1997. *Rabenvögel: Göttervögel - Galgenvögel. Ein Plädoyer im „Rabenvogelstreit“*. Leinfelden-Echterdingen: DRW-Verlag Weinbrenner.
- Erdösy, George. 1995. *The Indo-Aryans of Ancient South Asia: Language, Material Culture and Ethnicity*. Berlin: Walter de Gruyter.
- Finck, F. N. 1903. *Lehrbuch des Dialekts der deutschen Zigeuner*. Marburg: Elwert'sche Verlagsbuchhandlung.
- 1907 *Die Sprache der armenischen Zigeuner*. Sankt Petersburg: Imperial Science Academy.
- Fussman, Gérard. 1972. *Atlas linguistique des parlers Dardes et Kafirs. I. Cartes, II. Commentaire*. Paris: École Française d'Extrême-Orient.
- 1989. Gāndhārī écrite, Gāndhārī parlée. In: Colette Caillat (ed.), *Dialectes dans les literatures Indo-Aryennes*. Publications de L'Institut de Civilisation Indienne 55: 433–501.
- Gjerde, Lars (with Kristiansen, Knut). 1994. *“The Orange of Love” and Other Stories: The Rom-Gypsy language in Norway*. Oslo: Scandinavian University Press.
- Gjerdman, O. and E. Ljungberg. 1963. *The Language of the Swedish Coppersmith Gipsy Johan Dimitri Taikon: Grammar, Texts, Vocabulary, and English Word-Index*. Uppsala: Lundequist.
- Grierson, G. A. 1888. Doms, Jats, and the origin of the Gypsies. In: *Journal of the Gypsy Lore Society*, 1: 71–76.

- 1908. India and the Gypsies. In: *Journal of the Gypsy Lore Society*, New Series, 1 (1): 400.
- (Gen. ed.). 1919. *Linguistic Survey of India. Indo-Aryan family. North-Western group: VIII pt. I: Specimens of Sindhī and Lahndā*. Comp. and ed. by G. A. Grierson. 1919; pt. II: Specimens of the Dardic or Piśācha languages (including Kāshmiri).
- (Gen. ed.). 1922. *Linguistic Survey of India*. 11 vols: vol. 11, *Gipsy Languages*. Calcutta: Superintendent Government Printing, India.
- Hancock, Ian F. 1995. On the migration and affiliation of the Dōmba: Iranian words in Rom, Lom and Dom Gypsy. In: Matras, 1995: 25–51.
- 1998. *The Indian origin and westward migration of the Romani people*. (Manchaca, TX: International Romani Union). University of Austin, Texas.
- Halwachs, Dieter W. und Gerd Ambrosch unter Mitarbeit von Katharina Deman, Ursula Glaeser und Michael Wogg. 2002. *Wörterbuch des Burgenland-Romani (Roman)*. Arbeitsbericht 10 des Romani Projekts, Verein Roma/Oberwart: <http://romani.uni-graz.at/romani/publ/roman.en.shtml>
- Haugen, Einar. 1949. A note on the Romany “language”. In: *Norsk Tidsskrift for Sprogvidenskap* 7: 388–391.
- Hendriksen, Hans. 1986. *Himachali Studies; I. Vocabulary*. Det Kongelige Danske Videnskabernes Selskab, Historisk-filosofiske Meddelelser 48,1. Kommissionær: Munksgaard: København.
- Holzinger, Daniel. 1993. *Das Rómanes: Grammatik und Diskursanalyse der Sprache der Sinte*. Innsbruck: Verlag des Instituts für Sprachwissenschaft der Universität.

- Iversen, Ragnvald. 1944. *Secret Languages in Norway, volume 1: the Romany Language in Norway*. Norsk Videnskaps-Akademi, II Filosofisk-Historisk Klasse 1944, part 3. Oslo: Norsk Videnskaps-Akademi.
- Jettmar, Karl. 1975. *Die Religionen des Hindukusch*. Die Religionen der Menschheit; Bd. 4, 1. Stuttgart: Kohlhammer.
- Kaul, P. K. 2006. *Pahāri and Other Tribal Dialects of Jammu*. Two volumes. Delhi: Eastern Book Linkers.
- Kenrick, Donald. 2004. *Gypsies: From the Ganges to the Thames*. Hatfield: University of Hertfordshire Press.
- Krougly-Enke, Nikita. Preservation of the Nostratic Heritage and Renewal of Animal Names in Eskaleutian Languages:
[http://www.nostratic.ru/books/\(67\)krugly-enke1.pdf](http://www.nostratic.ru/books/(67)krugly-enke1.pdf)
- Kulikov, Leonid. 2007. Review of Southworth, in *Acta Orientalia Vilnensia*. 8/1, pp. 173–177.
- Kutlik-Garudo, Igor E. 1993. Sprachliches und Zeitliches zur Zigeunerimmigration in Armenien und Palästina im Mittelalter. In: *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 143: 265–286.
- Lindell, Lenny och Kenth Thorbjörnsson-Djerf. 2008. *Ordbok över svensk romani: Resandefolkets språk och sånger. Inledning, grammatik och bearbetning av Gerd Carling*. Stockholm: Podium.
- Lorimer, D. L. R. 1939. *The Dumaki Language*. Nijmegen: Dekker and Van de Vegt.
- Lubotsky, Alexander. The Indo-Iranian substratum. In: *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. Papers presented at an international symposium held at the Tvärminne Research Station of the University of Helsinki 8–10 January 1999*.

- (Mémoires de la Société Finno-ougrienne 242.) Chr. Carpelan, A. Parpola, P. Koskikallio (eds.). Helsinki 2001, pp. 301–317: <http://www.ieed.nl/lubotsky/pdf/Indo-Iranian%20substratum.pdf>
- Macalister, St. 1914. *The language of the Nawar or Zutt – The Nomad Smiths of Palestine*, JGLS – Monographs No. 3.
- Mallory, J. P. and D. Q. Adams. 2006. *The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World*. Oxford: Oxford University Press.
- Marcantonio, Angela (ed.). 2009. *The Indo-European Language Family: Questions about its Status*. Washington, DC: Journal of Indo-European Studies Monograph Series, No. 55.
- Masica, Colin. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Matras, Yaron (ed.). 1995. *Romani in Contact*. Amsterdam: Benjamins.
- 2002. *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.
- 2009. Defining the limits of grammatical borrowing. In: Marcantonio, Angela, 11: 1–25.
- Matras, Yaron, Peter Bakker and Hristo Kyuchukov (eds.). 1997. *The Typology and Dialectology of Romani*. Amsterdam: Benjamins.
- Metzger, Mathias. Etymological glossary of Indic words in Romani. Unpublished manuscript.
- Miklosich, F. 1872–1888. Über die Mundarten und die Wanderungen der Zigeuner Europas. In: *Denkschriften der philosophisch-historischen Classe der kaiserlichen Akademie der Wissenschaften*, Bd. 21–23, 25–27, 30–31.

- Morgenstierne, Georg. 1926. *Report on a Linguistic Mission to Afghanistan*. Oslo: Instituttet for Sammenlignende Kulturforskning. (Serie C I–2).
- 1927a. The linguistic classification of Dardic and Kafiri. *Oostersch Genootschap in Nederland* (Verslag van het 5de Congres), pp. 31–32.
- 1927b. *An Etymological Vocabulary of Pashto*. Oslo: Jacob Dybwad.
- 1929. *Parachi andOrmuri*. Oslo. (Indo-Iranian Frontier Languages 1).
- 1930. Notes on Torwali. In: *Acta Orientalia* 8: 294–310.
- 1932. *Report on a Linguistic Mission to North-Western India*. Oslo: Instituttet for Sammenlignende Kulturforskning Serie C III – 1.
- 1938. *Iranian Pamir Languages: Yidgha-Munji, Sanglechi-Ishkashmi and Wakhi*. Oslo (Indo-Iranian Frontier Languages II).
- 1940. Notes on Bashkarik. In: *Acta Orientalia* 18: 206–257.
- 1941. Notes on Phalura, an unknown Dardic language of Chitral. Oslo (*Skifter utgitt av det Norske Videnskaps Akademi i Oslo*, 1940, II. Historisk-Filosofisk Klasse, 5).
- 1942. Notes on Dameli, a Kafir-Dardic language of Chitral. In: *Norsk Tidsskrift for Sprogvidenskap* 12: 115–198.
- 1947. Metathesis of liquids in Dardic. In: *Festskrift til Prof. Olaf Broch på hans 80-Årsdag* (Skifter utgitt av det Norske Videnskaps Akademi i Oslo, II. Historisk-Filosofisk Klasse, 2), pp. 145–154. (Reprinted in Morgenstierne 1973: *Irano-Dardica*, pp. 231–240).

- 1950. Notes on Gawar-Bati. Oslo (*Skifter utgitt av det Norske Videnskaps Akademi i Oslo*, II. Historisk-Filosofisk Klasse, 1).
- 1954. The Waigali Language. Offprint from *Norsk Tidsskrift for Sprogvidenskap* 17: 146–324.
- 1956. *The Pashai Language*. 3. *Vocabulary*. Oslo. (Indo-Iranian Frontier Languages III).
- 1965. Notes on Kalasha. In: *Norsk Tidsskrift for Sprogvidenskap* 20: 183–238. (Reprinted in *Indo-Iranian frontier languages* 4 (1973a)).
- (without year specification) A note on the Khetrani dialect of Lahnda. In: *Acta Orientalia*, pp. 173–174.
- 1973a. *Indo-Iranian Frontier Languages*. Oslo: Universitetsforlaget (2d ed. rev. and with new material). (Serie B, Skifter – Institute for Comparative Research in Human Culture, Oslo, Vols. 11, 35, 40, 58). (Vol. 1: *Parachi andOrmuri*; Vol. 2: *Iranian Pamir Languages*; Vol. 3: *The Pashai Language* (1. Grammar, 2. Texts and Translations, 3. Vocabulary); Vol. 4: *The Kalasha Language*).
- 1973b. *Irano-Dardica*. Beiträge zur Iranistik, 5. Wiesbaden: Ludwig Reichert.
- 1974. *Etymological Vocabulary of the Shughni Group*. Beiträge zur Iranistik, 6. Wiesbaden: Dr. Ludwig Reichert.
- Pischel, Richard. 1883. Das Volk der Zigeuner. In: *Deutsche Rundschau* 36.
- 1894. Beiträge zur Kenntnis der deutschen Zigeuner. In: *Festschrift zur 200 jährigen Jubelfeier der Universität Halle*, Halle.
- Renfrew, Colin. 1987. *Archaeology and Language: The Puzzle of Indo-European Origins*. Cambridge: Cambridge University Press.

- Rozwadowski, J. 1936. *Wörterbuch des Dialekts von Zakopane*, Kraków (= Mémoires de la Commission Orientalistes No. 21).
- Salomon, Richard. 2002. Gandhari and the other Indo-Aryan languages in the light of newly-discovered Kharoṣṭhi manuscripts. In: N. Sims-Williams (ed.), *Indo-Iranian Languages and Peoples*. Oxford: The British Academy, pp. 119–134.
- Sampson, John. 1923. On the origin and early migrations of the Gypsies. In: *Journal of the Gypsy Lore Society*, Third Series, 2: 156–169.
- 1926/1968. *The Dialect of the Gypsies of Wales. Being the Older Form of British Romani Preserved in the Speech of the Clan of Abram Wood*. Oxford: Clarendon Press.
- 1927. Notes on Professor R. L. Turner's "The position of Romani in Indo-Aryan". In: *Journal of the Gypsy Lore Society*, Third Series, 6: 57–68.
- Schmidt, Ruth Laila and Razwal Kohistani. 2008a. *A Grammar of the Shina Language of Indus Kohistan*. Wiesbaden: Harrassowitz.
- Schmidt, Ruth Laila and Vijay Kumar Kaul. 2008b. A comparative analysis of Shina and Kashmiri vocabularies. In: *Acta Orientalia* 69: 231–301.
- Sharma, D. D. 1989. *Tribal Languages of Himachal Pradesh. Part One*. Delhi: Mittal Publications.
- 1990. *Tibeto-Himalayan languages of Uttarakhand 2: Studies in Tibeto-Himalayan Languages-3*. New Delhi: Mittal Publications.
- 1992. *Tribal languages of Himachal Pradesh. Part Two*. New Delhi: Mittal Publications.

- 1996. *Tribal Languages of Ladakh. Part One: A Concise Grammar and Dictionary of Brok-skad*. New Delhi: Mittal Publications.
- Southworth, Franklin C. 2005. *Linguistic Archaeology of South Asia*. London: Routledge Curzon.
- Stampe, David and Patricia Donegan. Austroasiatic Languages: Munda (Eastern India) and Mon-Khmer (NE India, mainland SE Asia, Malaysia, Nicobars):
<http://www.ling.hawaii.edu/austroasiatic/>
- Tálos, Endre. 1999. Etymologica Zingarica. In: *Acta Linguistica Hungarica* 46: 215–268.
- Thesleff, A. 1901. *Wörterbuch des Dialekts der Finnländischen Zigeuner*. Helsinki (= Acta Societatis Scintiarum Fennicae XXIX).
- Trail, Ronald L. and Gregory R. Cooper (compilers). 1999. *Kalasha Dictionary; With English and Urdu*. Islamabad: National Institute of Pakistan Studies and Summer Institute of Linguistics. (Studies in Languages of Northern Pakistan 7).
- Turner, Ralph Lilley. 1926. The Position of Romani in Indo-Aryan. In: *Journal of the Gypsy Lore Society* 5/4: 145–188.
- 1959. Transference of aspiration in European Gypsy. In: *Bulletin of the School of Oriental and African Studies*, University of London Vol. 22, No. 1/3 (1959), pp. 491–498.
- 1962–1966. *A Comparative Dictionary of the Indo-Aryan Languages*. London: Oxford University Press. Includes three supplements, published 1969–1985.
- Van den Broek, Roelof and Inez Wolf Seeger. 1972. *The Myth of the Phoenix according to Classical and Early Christian Traditions*. Leiden: E. J. Brill.

- Varma, Siddheshwar. 1938. The dialects of the Khaśālī group. In: *Journal of the Royal Asiatic Society of Bengal* 4: 1–65.
- 1948. *The Bhalesī Dialect*. Royal Asiatic Society of Bengal Monograph Series Vol. IV. Calcutta: Royal Asiatic Society of Bengal.
- Vekerdi, József. 1981. On the social prehistory of the Gypsies. In: *Acta Orientalia Academiae Scientiarum Hungaricae* 35: 243–254.
- 2006. An 18th-Century Transylvanian Gypsy vocabulary. In: *Acta Orientalia Academiae Scientiarum Hungaricae*. 59 (3): 347–360.
- Von Train, Josef K. 1832. *Chochemer Loschen: Wörterbuch der Gauner- und Diebs- vulgo Jenischen Sprache nach Criminal-acten und den vorzueglichsten Huelfsquellen fuer Justiz-, Polizei- und Mauthbeamte, Candidaten der Rechte, Gendarmerie, Landsgerichtsdienner und Gemeindevorsteher*. Regensburg: In Commission bei Friedrich Pustet.
- West, M. L. 2007. *Indo-European Poetry and Myth*. Oxford: Oxford University Press.
- Windfuhr, G. L. 1970. European Gypsies in Iran. A first report. In: *Anthropological Linguistics* 12: 271–292.
- Wolf, Siegmund Andreas. 1960. *Grosses Wörterbuch der Zigeunersprache (romani tsiw). Wortschatz deutscher und anderer europäischer Zigeunerdialekte*. Mannheim: Bibliographisches Institut.
- Woolner, Alfred C. 1913–1914. The Indian origin of the Gypsies in Europe. In: *Journal of the Panjab Historical Society* 2: 136–141.
- 1928. Aśoka and the Gypsies. In: *Journal of the Gypsy Lore Society*, Third Series, 7: 108–111.

- Zoller, Claus Peter. 2005. *A Grammar and Dictionary of Indus Kohistani. Volume 1: Dictionary*. Berlin, New York: Mouton de Gruyter.
- 2009. Genitive Marking of Subjects in West Pahārī. In: *Acta Orientalia* 2008: 69, 121–151.
- (forthcoming). *Bangani and Languages of North-Western South Asia: An Exploratory Study*.