Introduction

Peter Bakker
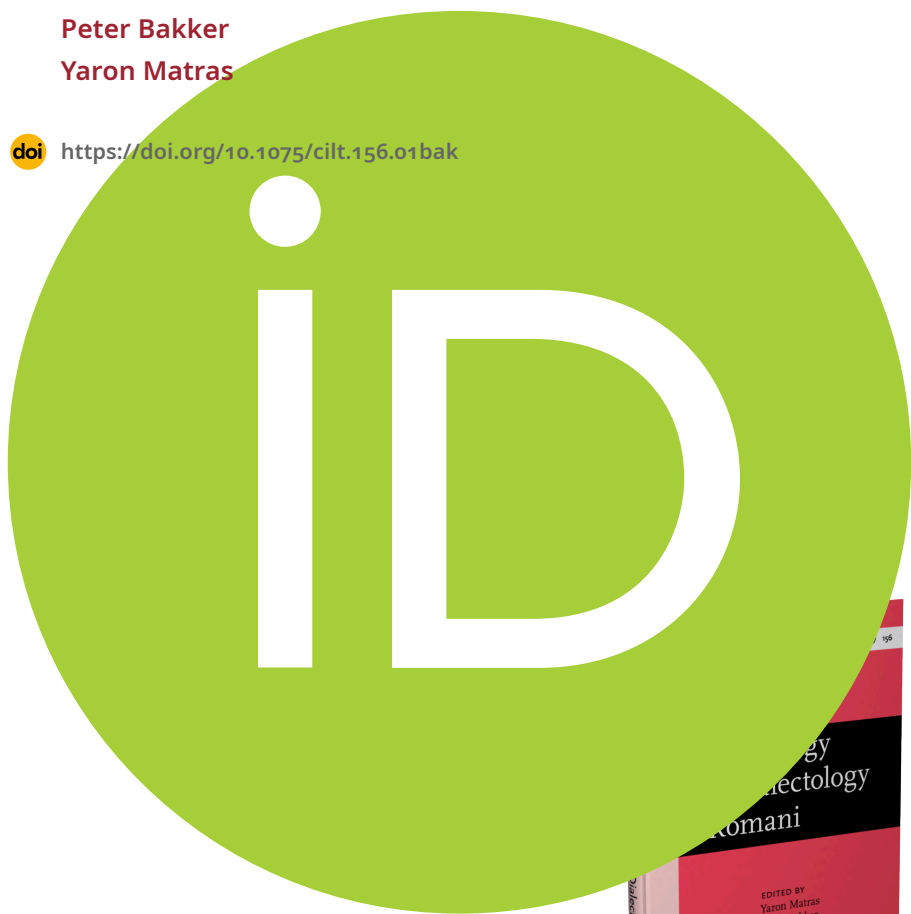
Yaron Matras

INTRODUCTION

PETER BAKKER
University of Aarhus
and
YARON MATRAS
University of Manchester

1. Romani linguistics: a very brief history

Romani is a minority language in all societies where it is spoken, but in terms of numbers of speakers Romani actually ranks among the more widely spoken languages of the world. With an estimated number of between five and ten million, it has more speakers than Catalan or Danish. This volume deals with two central topics in the study of the Romani language in which significant progress has been made in recent years: typology and dialectology. Recent descriptions of hitherto unknown or only sparsely documented dialects have broadened the database for both comparative and typological work on varieties of the language. Insights from this work have in turn begun to serve as input for more general typological studies, with Romani figuring both in larger language samples and in studies of single phenomena. The position of Romani as a *Balkanized Indic language* (Matras 1994a) and its hybrid patterns of morphosyntax are of special interest in this context.

Romani dialects were being documented as early as the sixteenth century. In 1542 Andrew Boorde published a short sample of British Romani in his *Fyrst Boke of the Introduction of Knowledge*, and the first Romani vocabulary, based on a northern continental dialect, was published by Bonaventura Vulcanius in 1597 (Kluge 1901:91, 113-114). An Italian theatre play from 1646 contains some dialogues in Romani (Piasere 1994); in 1668 the Ottoman writer Evliya Çelebi presented samples of Balkan Romani (Friedman & Dankoff 1991), and Job Ludolf followed in 1691 with a German Romani wordlist (Kluge 1901:172-173). It was not until the second half of the eighteenth century, however, that the grammar and the origin of the Romani language became the subject of inten-

sive scholarly discussion, leading ultimately to the wholesale dismissal of vari-
ous speculations that had flourished earlier (cf. Hancock 1988).

Between 1760 and 1785 various persons were engaged in Romani-related
research — among them William Marsden and Jacob Bryant in England, the
Hungarian Istvan Vali in the Netherlands, Christian Büttner, Johann Rüdiger,
and Heinrich Grellmann in Germany, and H.L.C. Bacmeister, Simon Pallas and
even Catherine the Great in Russia. They discovered, partly in cooperation and
partly independently of one another, similarities between Romani and living
Indo-Aryan languages. Grellmann is usually credited with the discovery of the
Indic origin of Romani, probably since he devoted to this topic an entire chapter
in his book on Gypsies (Grellmann 1787), a work which was subsequently
translated into English and other languages. While Grellmann not only made use
of his colleagues' materials, but also plagiarized other sources extensively (cf.
Ruch 1986), it was Rüdiger whose original essay on Romani (Rüdiger 1782) in
fact constituted the breakthrough.

Rüdiger was critical both of society's treatment of the Gypsies and of the
contemporary scholarly discussion concerning their language and origin.
Relying on wordlists collected mostly by Büttner, he first noticed lexical simi-
larities between Romani and Hindustani, which he went on to explore empiri-
cally. With a short description of the Hindustani language written by a mission-
ary in hand, Rüdiger sought the help of a native speaker, a Romani woman of
the Sinti group named Barbara Makelin (cf. Adelung 1815:30). He thus obtained
both Hindustani and Romani translations of phrases and paradigms and was able
to show that the two languages shared not only some basic vocabulary, but also
the essentials of inflectional morphology and some syntactic features. As a
typologist, Rüdiger's achievement is not only illustrating the shared origin of the
two languages, but also explaining the differences between them (cf. Matras
1994a:1; Haarmann 1989). He described the emergence of the definite article in
Romani, for example, as a case of the functional reanalysis of an Indic
demonstrative following the model of the contact language, and noted the heavy
impact of language contact in areas such as word order and indeclinable parti-
cles. Thus, as early as the late eighteenth century data on Romani was already
providing scholars with insights into the mechanisms of contact-related gram-
matical change and the principles of areal linguistics.

By the time August Pott had begun his voluminous work consisting of a
grammar and a comprehensive dictionary (Pott 1844-1845) several dozen pub-
lished sources on various Romani dialects had become available. It was Pott,
nowadays considered the founder of Romani philology, who concluded on the
basis of these sources that, despite variation and the strong influences of foreign
contact languages, Romani dialects clearly form a linguistic unity. Empirical

Romani dialectology on a large and comprehensive scale can be said to have started with the work of Franz Miklosich. His twelve-part opus on the *Speech Varieties and Migrations of the Gypsies of Europe* (Miklosich 1872-1880) is not only a comparative study of the most important Romani dialects, but also a meticulous historical and typological study of the language as a whole. On the basis of a comparative study of Indo-Aryan languages and their various stages of development, relying especially on evidence for the formation of the nominal paradigm, Miklosich concluded that the ancestors of the Roma must have left India around the tenth century, following the transition from Middle to New Indo-Aryan and the collapse of the ancient case inflection system. Miklosich also recognized various strata of lexical borrowings in the language, which he regarded as evidence for the migration route which the Gypsies had taken on their way to Europe and for the pathways of dialectal dispersion within Europe thereafter. Perhaps his most important observation pertained to the significance of the Greek component. Its presence in all dialects of Romani led Miklosich to conclude that a Greek-speaking area had been the European homeland of all Gypsies before their dispersion across the continent. It is only in the last decade that Miklosich's observations and their implications for an areal-typological classification of Romani have again become the focus of linguistic research and that Romani has been dealt with in the context of Balkan *Sprachbund* phenomena (see e.g. Friedman 1985; Matras 1994a and 1994b).

The first three decades of the twentieth century saw significant achievements in the description of single Romani dialects, the most outstanding one being John Sampson's grammar and etymological dictionary of the dialect of the Gypsies of Wales (Sampson 1926). A thriving discussion also emerged surrounding the origins of Romani in India (see de Goeje 1901; Sampson 1907; Turner 1926) as well as the connection between the European Gypsy dialects and the Indo-Aryan varieties spoken by the Dom and the Lom in the Near East and Armenia respectively (see Macalister 1909-1913; Finck 1907; also surveyed in Hancock 1988). Modern Romani-related research has since been mostly concerned with the following problems: 1) The unity and diversity of Romani dialects and their implications for both dialect classification and linguistic origins; 2) The impact of language contact on linguistic change, including grammatical borrowing and contact-induced internal innovation, as well as the retention of Romani vocabulary in instances of language shift; 3) The sociology of the language, in particular questions of status, codification, and standardization; and finally, more recently, 4) Romani in the context of current theoretical issues in general linguistics.

It would be beyond the scope of this Introduction to take an inventory even of the main studies published in the past few decades (but see Bakker & Matras,

forthcoming). Noteworthy however is the unprecedented upsurge of publica-
tions, especially monographs, during the 1990s. This latest period has seen ex-
tensive work in all four domains named above. As a representative sample of
dialect descriptions we mention Boretzky's (1993) study of the Balkan dialect
Bugurdži, Boretzky's (1994) outline of the Vlach dialect Kalderaš, Hancock's
(1995a) Vlach grammar, Igla's (1996) study of the Vlach variety of Ajia
Varvara, Halwachs' (1996) outline of the Central dialect called Roman as
spoken in the Burgenland, and Cech & Heinschink's (1997) description of the
Balkan Sepečides variety. Language contact surfaces in numerous case-studies,
and we note in particular Boretzky & Igla's surveys of loan morphology (1991)
and loan phonology (1993) as well as the collections by Bakker & Cortiade
(1991) and Matras (1995 and forthcoming). Advances in the codification of
Romani for both descriptive and applied purposes are represented by the
appearance of the dictionaries by Demeter (1990) for Kalderash Vlach, by
Hübschmannová et al. (1991) for the Central dialect of the Czech and Slovak
Republics, and by Boretzky & Igla (1994a) for the Vlach and Balkan varieties of
southeastern Yugoslavia. Finally, attempts to discuss Romani dialects in the
context of current theoretical frameworks are best represented by Holzinger's
(1993) study of Sinti as well as Matras' (1994a) study of Kelderaš/Lovari, both
applying discourse-based functionalist methodologies. But general linguistic
typology has also recently taken an interest in specific features of Romani, as
exemplified by Plank (1995; also Payne 1995), Van der Auwera (in press), and
others.

2. *Core typological features and the unity of Romani*

Despite over two centuries of research on the origin of Romani and its
European dialects, ignorance of the fact that Romani constitutes a distinct and
easily definable linguistic entity is still widespread. A superficial indication of
speakers' own awareness of linguistic unity is the term *romanés* (occasionally
also *rómnes*, *romanéh*, *romanó* or *róman*) which they use to designate their lan-
guage. In the light of all that has been written on Romani in the past two cen-
turies it is no longer necessary to argue the case for the structural unity of the
language; we choose nonetheless to scan some of the core structural features of
Romani in this section. This is intended primarily to facilitate access by non-
Romanologues to the more specific contributions to this volume, but may also
prove useful to those seeking additional confirmation of a well-defined, struc-
tural linguistic entity.

While diversity is doubtless most extreme at the lexical level, Romani dialects share a basic lexical core comprising the most frequently occurring items. This core is itself composed of various historical layers, showing some 700 Indic roots, around 70 Iranian, 40 Armenian, and perhaps 230 Greek items (cf. Boretzky & Igla 1994a). This composition is constant, not random, and is subject to little variation, and so it is taken to reflect an inherent shared origin rather than superposed convergence. But beyond core vocabulary, which in Romani is admittedly small, the language undoubtedly constitutes a grammatical unity.

Romani morphology is best characterized by a delicate balance of inflective features inherited from Older Indo-Aryan, agglutinative features which parallel some of the later developments in Modern Indo-Aryan and in Modern Indo-Iranian as a whole, and a more recent tendency towards analytic formation characterized by structural renewal and the grammaticalization of items of Indo-Aryan stock. These processes which involve inherited Indo-Aryan morphology partly overlap with grammatical borrowing of unbound, semi-bound, and in some cases even bound morphemes from the European contact languages. On the whole, different morphological formations may be found within the same grammatical categories, though some areas of Romani grammar may typically resort to specific patterns (for an extensive discussion see Elšik's contribution to this volume).

Romani has two genders (masculine and feminine), and two numbers (singular and plural). The expression of gender can be regarded as an inflective feature of nominal declensions, though a degree of agglutination is apparent in the regularity of the suffixes -o for masculine and -i for the feminine in the nominative singular of nouns, adjectives, and demonstratives, in participles including the active participles often used for past-tense formation of unaccusative verbs (predominantly expressing change of state; cf. Matras 1995b), in subject clitics, and reflected finally in the nominative definite articles o and i (alternating with e in some dialects) respectively. Number partly shares this regularity, with -e marking the plural in some nominal declensions, in adjectives and demonstratives, in participles, clitics, and in the definite article of some dialects. Number and gender agreement thus appear across all these categories, though there are nominal and adjectival declensions which do not take overt gender marking, and the rules for gender marking are typically different in thematic and athematic paradigms (see below). Agreement also extends to the genitive case, which agrees with the head noun of a possessive construction through 'Suffixaufnahme' (see Plank 1995:11-13; Payne 1995:288-289).

Internal inflection stands out in the system of demonstratives, which is typically quadripartite, encoding the source of knowledge about a referent (context vs. situation) as well as its accessibility or 'discreteness' (Matras, in press).

Derivation of both nouns and verbs however typically involves agglutinative suffixing procedures. Thus *-ni* forms feminine animates, *-to* ordinal numerals, etc.. In addition, the genitive case is used for derivational purposes in some dialects. Compounding is rare, and the use of verbal roots such as *ker-* 'do' or *d-* 'give' in complex formations can be regarded as a case for grammaticalization rather than as lexical compounding. There are a number of productive category-changing suffixes such as *-alo* and *-ano*, which derive adjectives from nouns, or *-ipen/-ibe* and variants, which derive nouns from adjectives and verbs. With verbs, agglutination is particularly outstanding in the formation of the causative by means of the infixes/suffixes *-av-*, *-ar-*, *-ker-* etc. (see contributions by Hübschmannová & Bubeník and by Cech & Heinschink to this volume), while passivization on the other hand involves inflective procedures including a change in the root of the verb, often with palatalization accompanied by a shift in stress (*dikh-él* 'sees' > *díčh-ol* 'is seen'). On the whole, the regular patterns allowing for causativization and passivization lend themselves to an interpretation of transitivization/de-transitivization as the central axis underlying verb derivational processes in the language.

Romani verbs are inflected for person/number by two series of inflective personal endings, correlating with the presence vs. absence of infixed extensions of the verbal stem. These extensions (*-d-*, *-l-*, *-t-*) are sensitive to phonological environment and show some dialectal variation. They have been interpreted as 'aorist' suffixes (Hancock 1995a), or alternatively, regarded as expressions of aspect rather than tense (see Holzinger 1993; Matras 1994a). Tense in the strict sense is expressed in most dialects by an agglutinating suffix *-as* which is external to aspect and — a typological rarity — to person marking. Renewal and consequent dialect variation is typical of the formation of the future tense. This is also an area where in some dialects analytic formation in the verb system may be encountered, with auxiliaries such as *l-* (lit. 'take') or particles such as *ka* (derived from *kam-* 'want' modelled on the Balkan future tense). Analytic formation may also encompass other areas of the verb in some dialects. Thus an analytic passive, based on the use of the copula as an auxiliary with a passive participle, is in use alongside a middle voice, based on the inflected passive (de-transitive) followed by the reflexive pronoun *pe(s)*. In a number of dialects, some unaccusative verbs will employ the copula auxiliary and a past participle rather than finite forms of the main verb in both past and present tenses.

Nowhere is the interplay of inflective, agglutinating, and analytic features as clearly represented as in the formation of nominal case (see the contribution by Matras to this volume). Romani preserves traces of the Old Indo-Aryan noun inflection in the nominative, the accusative, and the vocative. The accusative,

which is based on the Old Indo-Aryan genitive, is the most productive non-nominative case and is sometimes designated 'oblique', for two main reasons: Firstly, it forms the basis for further case formations. Secondly, it co-occurs with the oblique case of adjectives, adjectival demonstratives, possessives, and definite articles, where a nominative/non-nominative two-way distinction prevails. The accusative figuring as an oblique hosts a series of agglutinative suffixes expressing the dative, locative, genitive, instrumental/sociative, and ablative cases. While the inflective cases encode gender and number (see above) and occasionally declension class (though differentiation here is minimal, compared to Old Indic, Latin, or contemporary Slavic), the agglutinative markers are isomorphic and highly regular; they vary only to the extent to which they are subjected to partial phonological assimilation with the preceding oblique ending, rendering alternate forms such as locative -te/-de reminiscent of case agglutination in Turkic or other language groups. Analytic case formation with prepositions (usually gramaticalized Indic adverbs of location) generally co-exists with the inflective/agglutinative pattern, but often substitutes for the latter. Indeed, for the language as a whole analytic prepositions may be said to be gaining ground at the cost of case endings.

Suppletion is on the whole rare in Romani. It is best represented in the system of third person pronouns (though not pronominal clitics), where nominative stems differ from non-nominative (oblique) stems, a differentiation which in some dialects is mirrored by the system of definite articles as well. In addition, suppletion is often found in the present/future and indicative/non-indicative formation of the copula (see Boretzky's contribution to this volume) and in the present/past of the verb 'go' (dža-/g-). To those one might add the negated copula in some dialects, as well as verb negation on the whole, which often shows indicative/non-indicative suppletion. A weaker case for 'semi-suppletion' in contemporary Romani could be made for reflections of the Old Indic present/past stem alternations which are still present in some verbs such as sov-/su- 'sleep', mer-/mu- 'die', per-/p- 'fall'.

Morphosyntax is more strongly influenced by contact phenomena than morphology proper, though some Indic features are preserved, in particular the non-nominative possessor, usually marked by the accusative/oblique (rather than the dative) in a possessive construction with the copula. Word order within the noun phrase only partly resembles Indo-Aryan patterns. For adjectives and demonstratives the dominant order is Det-N, with N-Det as an occasional though highly marked variant. Numerals always precede the noun, as do genitive attributes (though N-Gen may appear in lexical compositions). The definite and indefinite articles, a Romani innovation, precede the noun and thus conform to the earliest significant European contact language, Greek. Word order in the

clause is often subject to variation and innovation through contact influences (see Boretzky 1996b), though in overall basic terms Romani word order can be classified as Balkan. Thus, thematic elements precede the verb while rhematic elements follow. This leads to an alternation of SV and VS as 'establishing' and 'connected' respectively (Matras 1995c). In addition, SV prevails in indicative complements while VS prevails in non-indicative complements. Romani has a striking tendency towards finiteness. Clause combining is consequently carried out almost entirely by unbound conjunctions. Subordinating conjunctions are partly derived from interrogatives or deictics, and partly borrowed. There is a strong tendency to borrow adversatives and sentential particles from contemporary contact languages (Matras 1996).

The most salient syntactic Balkanisms in Romani are the lack of an infinitive, the use of separate conjunctions with factual and non-factual complementation, the spread in some dialects of object doubling (see Bubeník's contribution to this volume), the use in some dialects of a future particle derived from the verb 'want', and perhaps some of the rules of lexical case assignment, the emergence of a middle voice, as well as the development of past tense evidentiality in some unaccusative verbs (Matras 1995b). Some of these features have been reversed or altered in some dialects, though traces of the Balkanized core structure remain. For example, a 'new infinitive' (Boretzky 1996a) is introduced in some Romani dialects based on the finite form of the third person singular present, though it is still introduced by the non-factual complementizer *te*.

Language contact plays an important role both in specific dialects of Romani and with regard to the character of the language as a whole. All varieties of Romani have distinct inflectional and partly derivational paradigms for vocabulary (nouns, verbs, adjectives) pre-dating the contact with Greek and for that acquired thereafter, labelled 'thematic' and 'athematic' repectively (see Bakker's contribution to this volume; Hancock 1995a; cf. also Boretzky 1989). While the thematic paradigms are Indic in origin, the athematic ones are typically a combination of Indic inflection and derivation with borrowed affixes of apparently Greek origin (Bakker, this volume), though other contact languages may occasionally contribute to the athematic morphology of single dialects as well, as in the case of plural endings on the noun. Borrowing of linguistic material is strongest in the area of unbound sentential and connective particles, including for example phasal adverbs, indefinites pronouns, and adversatives. Dialects affected more intensely by a particular contact language may show borrowed semi-bound items, such as prepositions or verbal particles from German in Sinti, or even bound morphemes, as in the case of Slavic aspectual markers and conditionals in some Northern, Central and even Balkan dialects, Slavic evidentials in some Balkan dialects, and Rumanian plural endings in Vlach.

The presence of significant layers of lexical and grammatical loans in Romani and the re-occurring tendencies of Romani dialects to undergo at least partial syntactic convergence with their co-territorial contact languages have led some researchers — mainly social scientists (for example Okely 1983; Willems 1995; Lucassen 1996), but recently also linguists (Wexler, forthcoming) — to express scepticism with respect to the existence of Romani as a coherent linguistic entity of Indic origin. Instead, the Indic component is seen as part of an Asian vocabulary collected along trading routes and markets, in other words as a trading slang, functionally akin, perhaps, to the secret vocabularies of itinerant groups in Europe (Cant, Shelta, Rotwelsch, Jenisch, Bargoens etc.) or to a pidgin. However, Romani comparative dialectology and a typological approach to Romani (analyzing it in the light of what is known about universals of language) provide us with tools to assess contact in its real dimensions and true impact on the language.

Romani as a unity is firstly characterized by a hybrid structure, owing its inflectional and derivational morphology and some patterns of morphosyntax to Indo-Aryan, and its sentence-level syntactic organization to the convergent development it underwent as part of the Balkan area during its early European phase (see Matras 1994a). Thus, Indic and non-Indic elements in the core structure that characterizes all varieties of Romani are not randomly distributed, but reflect the tendencies of languages in general to follow categorial differentiation in processes of language change and structural renewal. Moreover, Indo-Aryan morphology in Romani is not simply a fossilized remnant accompanying a random heap of Asian vocabulary, as suggested. Rather, its productivity can be seen at two levels: First, Indo-Aryan morphology is incorporated even into the inflectional paradigms applied to 'new words' (athematic elements). The Indic-derived markers for case, person, aspect, and tense are in fact equally used for thematic and athematic elements, and the systematic differentiation is reflected merely in their mode of attachment to the respective root. Second, and of no less importance, inherited Indic morphology constitutes a reservoire for creative processes of grammaticalization which accompany both internal and contact-motivated syntactic renewal. Thus the great majority of prepositions derive from Indo-Aryan adverbials (and have cognates among postpositions in Modern Subcontinental Indo-Aryan); adverbial conjunctions are recruited from among the class of Indic-derived interrogatives; Balkan complementational configuration is enabled through reanalysis of an Indic correlative particle as a non-factual conjunction (see Matras 1994a); the Balkan-type future particle is a de-lexicalized Indic-derived verb; etc. etc..

Not only is the distribution of the Indo-Aryan grammatical component very far from being random (or indeed marginal), but borrowing of actual grammati-

cal elements is also regular. Romani may be unique in its thematic/athematic morphological split, but when it comes to incorporating grammatical items from contact languages it follows the universals of grammatical borrowing attested in numerous contact settings (cf. Moravcsik 1978; Thomason & Kaufman 1988; also Harris & Campbell 1995): Superficial contact will result in occasional borrowing of unbound particles; dialects with more intense contact show borrowing of semi-bound elements and some phonemes; prolonged contact will lead to the adoption of some bound morphemes, and perhaps to some changes in word order, though not to a radical change in typological structure. It is clear, when comparing Romani dialects, that the major change in typological structure which distinguishes Romani as whole from Modern Subcontinental Indo-Aryan can be attributed to the process of language convergence in the Balkans and so to a prolonged and extremely intense contact with Greek in particular; in this respect, Miklosich's claim made more than a century ago of a Greek-speaking territory as the European homeland of the Gypsies can still be maintained on the grounds of contemporary methodology. (We return to the issue of language contact and its reflection in the present volume in section 4 below).

3. Dialectal diversity in Romani

Two language groups are not considered in the present context of variation within Romani. First, the so-called Para-Romani languages or dialects. The term relates to the preservation of a Romani-derived vocabulary within a non-Romani grammatical framework in Gypsy communities where language shift has taken place (see Boretzky & Igla 1994b; Bakker & Cortiade 1991; Matras, forthcoming). Para-Romani varieties are relevant to a discussion of dialect classification insofar as their lexical and partly phonological features allow some insights into earlier connections among the underlying dialects of Romani proper. English Para-Romani, also known as Angloromani, plays a marginal role in Hancock's contribution to this volume.

The other grouping that is excluded from this discussion are the Indic languages of Gypsies outside Europe, with the exception of course of the Romani dialects of out-migrants in Iran (see Windfuhr 1970; Djonedi 1996) or the Americas (e.g. Hancock 1976), specifically Near Eastern Gypsy or Náwari (also Nuri or Domari; see Macalister 1909-1913). Despite a striking affinity with Romani, Náwari is also quite distinct, both in parts of its Indic-derived morphology as well as in the impact of subsequent changes and contact layers (cf. Hancock 1995b). Comparisons of Romani with Náwari are drawn in this volume by Boretzky for the copula and by Bubeník for object clitics.

A detailed framework for Romani dialect classification is still lacking, and no comprehensive discussion of documented dialects is yet available. We will therefore restrict ourselves here to naming the major dialect branches, following the classification common in most recent studies and also applied in this volume, and to pointing out some of the types of isoglosses encountered. The reader is also referred to the list of dialect designations mentioned in the volume in the appendix to this Introduction, as well as to the Index of Subjects, where dialect names are included.

Romani can be divided into four main groups called Balkan, Vlach, Central, and Northern. The Northern branch constitutes a somewhat weaker entity, and its sub-branching into British Romani (now virtually extinct), Finnish Romani, Sinti-Manuš (German Romani), Polish-North Russian-Baltic, and sometimes Iberian Romani (now extinct) is often preferred. For all major dialect groups the geographical labelling takes into account the constitution of separate groups between the fifteenth and nineteenth centuries. Subsequent migrations have since resulted in various shifts in distribution as well as in interdialectal contacts and possibly in the emergence of new sub-divisions.

The Balkan dialects are spoken in the central and southern Balkan regions, mostly in Bulgaria, former Yugoslavia, Albania, Turkey, and Greece. They are sometimes subdivided into northern and southern Balkan. The subgroups often have names derived from Turkish, such as *Erli* or *Arli*, from Turkish *yerli* 'settled' (hence locally settled Roma), *Sepetči* (Turkish for 'basket-weaver'), or *Burgudži* (Turkish for 'drill-maker'). Muslim Roma in the Balkans often refer to themselves collectively as *Xoraxane* (or *Koraxane*) *Roma* or 'Turkish Gypsies', from *Koraxaj/Xoraxaj* 'Turkey', apparently a term used as a designation for non-Gypsies (cf. *ḳorāxāy* 'foreigner' in the Romani dialect called Romano of eastern Iran, Djonedi 1996; also Sampson 1926:183-184). Orthodox Roma living among Slavic-speaking peoples often refer to themselves collectively as *Dasikane Roma*. Balkan dialects have recently been promoted through publications and other institutional use of Romani especially in Bulgaria and Macedonia (see Friedman's contribution to this volume). Outlines of Balkan dialects are presented in this volume by Cech & Heinschink and by Igla.

The Vlach (also Vlax) dialects emerged in the Rumanian principalities probably during the period of Romani serfdom and slavery, which ended in the second half of the nineteenth century (see Hancock 1987). The distinction between 'Vlach' and 'non-Vlach' was introduced by Gilliat -Smith (1915-1916) as a classification of dialects in northern Bulgaria, but 'Vlach' has been used since to denote those dialects which share a number of structural features, the most obvious being the impact of Rumanian on vocabulary and loan-morphology. Subgroups usually have names derived from Rumanian or Hungarian occupational terms,

such as *Kalderaš* (also *Kelderaš*) from Rumanian *căldărar* 'coppersmith', Čurari from Rumanian *ciurar* 'sieve-maker', or *Lovari* from Hungarian for 'horse-dealer'. Vlach dialects are particularly widespread due to the outward migration of Roma following the abolition of serfdom from 1864 onwards, and the largest Romani population of any one country — that of Rumania — is almost entirely Vlach-speaking.

The Central dialects are spoken in the Czech and Slovak Republics, in southern Poland, by a minority of non-Vlach Romani-speakers in Hungary whose dialect is called *Romungro* or 'Hungarian Romani', and by a small group in the Austrian Burgenland and neighboring Slovenia. The Slovak dialect (spoken since the 1950s in the Czech part of Czechoslovakia as well) is represented in a relatively large number of Romani-language publications and provides the corpus for Elšik's contribution to this volume and, along with Hungarian Romani, for that by Hübschmannová & Bubeník.

The Northern dialects are, as mentioned above, probably the most diverse and widespread geographically. They include dialects as far apart as those of Wales and Finland. All documented western European Para-Romani varieties show in their vocabularies traces of this northern group, which is believed to have spread, subsequent to migrations from the Balkans, from southern Germany outwards during the fifteenth century. The group includes the dialects of the *Polska Roma* in Poland and the closely related Baltic and North Russian varieties, which still show traces of German influence in their vocabulary. A distinct entity consists of German Romani or the dialects of the groups whose self-designations are *Sinti* (in Germany, the Netherlands, and parts of Italy), *Manuš* (mostly in Alsace), or *Lalere* (formerly in Bohemia), all showing heavy impact of German on lexicon, morphosyntax, and phonology.

A superficial feature of dialect groups is the self-designation used by speakers. In the Northern dialects, the word *rom* plural *roma* is not used as an ethnonym, as in other dialects, but means only 'man/husband'. Instead, various ethnonyms are used in the Northern dialects including *Kalo* lit. 'black', *Manuš* lit. 'person', *Sinto*, *Romaničel*. On the other hand, only in the Vlach and Balkan dialects do subgroups use names denoting occupations.

There are only few isoglosses in Romani which might possibly reflect underlying subcontinental Indic isoglosses. The alternation of *s-* and *h-* in the copula (Vlach *som* 'I am', Sinti *hom*) is paralleled by the selection of different Middle Indo-Aryan verbs for the copula in subcontinental Indo-Aryan, rendering *s-* alongside *h-* paradigms (cf. contribution by Boretzky to this volume). *s-/h-* in Romani is on the other hand also a regular morpho-phonological alternation affecting person suffixes (*manges/mangeh* 'you want'), interrogatives (*sar/har* 'how'), case (*lesa/leha* 'with him'), and more. Sinti and the Central dialects are

considered *h*-dialects, though *h*-forms infiltrate other varieties as well. The inventory of deictic expressions varies somewhat, with the Indo-Aryan stems in *a-/o-, k-, d-*, and *v-* appearing in different combinations, though *(a)dava/(o)dova* for 'this/that' and *adaj/odoj* for 'here/there' are characteristically non-Vlach. The genitive suffix is *-kero* in non-Vlach and *-ko* in Vlach, again mirroring a subcontinental isogloss (Maithili *-ker*, Hindi *-kā*), though phonological simplification on European territory is possible and perhaps even more likely.

Thus in general, dialect differentiation can be taken to represent the result of phonological developments that must have occurred during the past four to five centuries or at any rate following the Europeanization of Romani. These may, as in the case of the *s-/h-* alternation, affect entire morphological paradigms and so constitute superficial grammatical differences. The palatalization of the past infix *-d-/-l-* for example is a general process which is carried out to varying degrees in different dialects, rendering regular paradigm differences as in *(a)rakh-l-om, rakh-lj-om, rakh-č-om, rakh-j-om* 'I found', *phen-d-om, phen-dj-om, phen-gj-um, phen-om* 'I said'. Similarly, there is variation in the copula and past tense personal endings of the first singular (*-om, -em, -um*) and second singular (*-an, -al*) as well as the past tense third singular (*-as, -a*). Interdialectal contact occasionally gives rise to 'contaminated' forms; thus the Central dialect has *sal* for 'you are', cf. neighboring Sinti *hal* and Vlach *san*. Some morphophonological developments of this type must therefore be regarded as secondary to the 'generic' division of dialects outlined above, and so as geographically superposed on them.

More 'generic' features include the distribution of morpheme stems (e.g. as above, for copula formation and deixis) and lexical roots (e.g. Vlach *kor*, Northern, Central, and Balkan *men* 'throat'; Vlach *punro* and Central *pindro*, Northern and Balkan *her* 'leg'). Some derivational formations are more productive in certain dialect groups: the Northern branch typically makes extensive use of genitives as well as of nominalization suffixes (*-ben*) for lexical derivation, while causativization is especially productive in Central and some Balkan dialects (see discussion in Hübschmannová & Bubeník and in Igla, both this volume). In some cases the phonological formation of the root may be said to be 'generic' (cf. Northern and Central *graj*, Vlach and Balkan *gras(t)* 'horse'; Northern, Central, and Balkan *maro*, Vlach *manro/marno* 'bread'). In others, dialect groups appear to have undergone separate secondary developments. Naturally, a complete inventory of dialect-specific sound changes would by far exceed the scope of this section. Some outstanding developments include jotation of initial vowels in the continental Northern dialects (*jamen* 'we', *jaro* 'egg'; cf. Vlach, Balkan, and Central *amen, ar(n)o*), the reduction of affricates to sibilants in Vlach (*šavo* 'boy', *ža-* 'go', Northern, Balkan, and Central *čhavo*,

dža-), secondary palatalization in Balkan and Vlach dialects spoken in some areas (*kher > ćher* 'house') with consequences for the case paradigm (*man-ge, tu-ke > man-dźe, tu-će* 'to me, to you'), and a shift to initial stress and subsequent syllable reduction is characteristic of Sinti (*krik* 'opposite', cf. Vlach *karing*). Some Balkan and Vlach dialects, most notably Kalderaš, retain an opposition [ʀːr], the uvular reflecting an underlying Indo-Aryan retroflex dental.

To summarize the internal features distinguishing Romani dialects and dialect groups, we may divide isoglosses into 1) distribution of stems, 2) the extent and distribution of a closed set of phonological developments affecting mophological paradigms, 3) productiveness of certain derivational procedures, 4) stem-internal phonological changes, and 5) secondary, geographically restricted phonological changes. Classes 1, 3, and 4 tend to be 'generic' features of the underlying dialect groups, class 2 includes large-scale geographical phenomena, while class 5 accommodates more local phenomena. Alongside these classes of internal features, there are numerous external features deriving from contact with various languages. As stated, a full inventory is not possible here, but we mention again the extensive borrowing of unbound and indeclinable morphemes (sentential and connective particles, phasal adverbs, etc.), the replication of the Slavic system of verbal aspect (*šunel* '(he) hears)', *pošunel* 'he listens') in the Polish-North Russian varieties and the Central dialects, and the use of German prepositions and verbal particles in Sinti.

4. This volume

The present collection reflects recent developments in the area of typology and dialectology[1]; Romani being a language with only a very recent written tradition and fragmentary documentation in past centuries, any attempt at interpreting Romani data in an historical depth relies on the comparative study of dialects and structural features. At the same time, the contributions attempt to point out features of Romani which are likely to be of interest to general linguists in the context of studies of universal properties of language.

Peter Bakker's paper is a study of a phenomenon which is rather rare in the languages of the world, but which is present in all Romani dialects to a greater or lesser extent: the paradigmatic-inflectional dichotomy between 'old' and 'new' words. Old words are mostly words from Asian languages (including Byzantine Greek), new words are from European languages. Bakker argues for a Greek origin of the patterns of morphological adaptation of loanwords, although, strangely enough, these patterns are not used for the bulk of the Greek lexical component in Romani.

Victor Elšik's paper focuses on a particular dialect of Romani, a variety spoken mostly in the Czech and Slovak Republics, but draws general comparisons with other dialects as well as with the related Modern Subcontinental Indo-Aryan languages. Elšik's typological classification of Romani is thus of general validity for the language as a whole. Working within the framework of Skalička's Prague School Typology, Elšik defines Romani as a typologically mixed language in almost all aspects of its grammar, showing agglutinative, inflective, and analytic features, though current tendencies show a drift towards stronger analyticity.

Yaron Matras' paper provides a case study which confirms this last point: the drift away from inflectional nominal case-marking and on to an analytic system of prepositions. Matras compares data from all main branches of Romani and postulates three hierarchies which allow to explain (and predict) the structural behavior of a given case form in a given variety. Semantic and pragmatic features play a role here. Some can be related to universal hierarchies and confirm observations made elsewhere, though some particular traits are also apparent.

Vít Bubeník deals with typological aspects of Romani morphosyntax. Drawing on categories developed in the context of Prague School Typology and Dik's Functional Grammar, Bubeník discusses object doubling as a syntactic phenomenon which originates in the grammaticalization of pragmatic functions of noun phrases, and draws a comparison with related structures in the Balkan languages. Areal factors, it is thus argued, play a role in motivating this development, but universal tendencies are involved as well. The question of the trigger behind the Romani structure is further complicated by the existence of a similar phenomenon in Náwari, the Palestinian Gypsy language, though here its appearance could be related to contact with Arabic and so to areal factors, too.

Norbert Boretzky's discussion of the development of the Romani copula is an example of how historical reconstruction in Romani draws on comparative dialectology. Boretzky makes use of the large corpus of data on Romani dialects of both published written sources and recent dialectological fieldwork. Indeed, his endeavor might not have been possible as recently as five years ago due to the little data available at the time. Boretzky posits different roots as the sources for the different verb stems, illustrating the principle of suppletion which Elšik identifies as a feature of inflectional languages. He also discusses the grammaticalization of the copula, its use as an inchoative suffix and a past tense suffix, and the extension of the verb 'come' to mean 'become'.

Milena Hübschmannová and Vít Bubeník also combine descriptive and historical work in their analysis of derivational verbal morphology in Romani. Based on fieldwork undertaken in the Czech and Slovak Republics, they deal

with phenomena which are rather a typological oddity in the languages of Europe. Romani is one of the few languages in Europe with regular morphologically formed causatives, passives, reflexives, and inchoatives. The paper also mentions a semantic shift from causative to iterative in Romani, a change also reported from other languages but for which no explanation has yet been proposed. Hübschmannová & Bubeník further show that at least some Romani dialects (namely Hungarian and Slovak Romani) have second causatives. There are clear formal parallels for the phenomenon in older stages of the Indo-Aryan languages, from which the Romani structures in question can be derived. But although causativization is an inherited phenomenon, language contact is seen to play a role in the preservation of causatives, which are more productive in dialects which share the feature of morphological causativization with their coterritorial language; this point is reinforced by Cech & Heinschink in their outline, in this volume, of the Sepečides variety.

Birgit Igla describes a number of properties of Romani as spoken in some communities in the Rhodopes, a mountain range in southwestern Bulgaria. The dialect shows a number of unique features. Case endings appear alongside both prepositions and postpositions. In addition, there are double prepositions, and prepositions which govern the accusative case (whereas other dialects show nominative). Rhodopes Romani is lexically conservative and preserves a number of Indic words which are otherwise hardly known in Bulgarian or Balkan dialects of Romani. It borrows aspectual prefixes from Slavic, a phenomenon also reported from other dialects in close contact with Slavic languages. Some particular phonological developments, which have re-shaped the form of inflectional paradigms, may be the result of phonological convergence with local Bulgarian dialects.

Petra Cech and Mozes Heinschink provide a general outline of the Romani dialect of the Basket-weavers (Sepečides) of Izmir in Western Turkey, a variety now endangered. The dialect is heavily influenced by Turkish in the areas of syntax, vocabulary and idioms, and phonology. On the other hand, the variety is extremely conservative. The verb *therava* 'I have', for instance, which is very rare in Balkan varieties of Romani, is still used here. Particularly striking in this dialect is the number of different causative constructions. The paper reiterates Hübschmannová and Bubeník's conclusion that morphological causatives are especially preserved in areas where the contact language shows similar constructions.

The last two papers deal with written Romani, albeit with two quite distinct aspects of it. Victor Friedman analyzes recent publications in Romani in the Republic of Macedonia, whereas Ian Hancock discusses the Romani language in the novels of the Victorian author George Borrow.

Victor Friedman reported in an earlier paper on the Standardization Confer-
ence in Macedonia and its likely implications (Friedman 1995). Romani has en-
joyed official recognition in Macedonia for a number of years now, and it is
used both in broadcasting and in print. In his present contribution Friedman fo-
cuses on features of a bilingual Romani/ Macedonian magazine. The coexistence
of several Romani dialects in the same country motivates authors to alternate
between single-dialect preferences and dialect mixing. Some of the choices relate
to orthography and phonology, others to morphological and syntactic features.
Preference is often given to the Arli dialect, which is both widespread and pres-
tigious among the Roma of Macedonia. It is especially noteworthy that the pub-
lication appears to follow the proposal adopted at the 1992 Standardization
Conference. On the other hand, despite the use of a local dialect the magazine's
international orientation is reflected in its choice of the Latin alphabet.

Ian Hancock discusses the individual 'dialect' of a non-Rom who has had
an enormous impact on both popular and scientific appreciation of the Romani
language during the nineteenth century and thereafter, even to this very day.
George Borrow wrote a number of books on the Gypsies of Britain and Spain
and had a keen interest in their language. He included in his novels Romani vo-
cabularies, phrases, and texts acquired through personal contacts with
Romanichals and Gitanos. Hancock shows that this material is unreliable from a
dialectological viewpoint. Borrow was on the one hand fully aware of the unity
of the Romani language; he met, and conversed with, Gypsies in Spain, Britain,
Russia, Hungary and other places. But his appreciation of this unity allowed
him to mix dialectal forms and structures rather freely, which makes his linguis-
tic surveys rather questionable sources.

We return at this point to the issue of contact and its reflection in Romani,
and to the point made above that contact is relevant to an appreciation of both the
typological features and the dialectal diversity of Romani. Single phenomena
discussed by the authors provide several illustrations of this. Bakker shows how
contact with Greek led to the adoption of a Greek strategy for loanword incorpo-
ration, and how this strategy has since been applied in Romani to cope with
ever-changing strata of loan vocabulary. Matras illustrates how the shift to
prepositions and the drift towards more analytic case formation relates to the
overall re-structuring of the language, particularly in the context of
Europeanization and Balkanization. Bubeník considers the case for
Balkanization in relation to object doubling. And while causativization in
Romani is clearly Indic-derived, it is argued both by Hübschmannová &
Bubeník and by Cech & Heinschink that contact with 'morphologically
causativizing' languages reinforces the productiveness of these structures. Igla
considers the possibility that the unusual vowel shifts in the Rhodopes dialect

mirror sound changes in the local Bulgarian dialect. Finally, Friedman discusses some of the problems facing the Romani written press in an area where several Romani dialects coexist, where the technical facilities for the production of written texts can only be accessed via the national (contact) language, but where writing in Romani is partly, if not significantly motivated by contact with international (i.e. non-Macedonian) varieties of Romani.

So is there a particular insight which Romani dialectology can offer towards a deeper understanding of universals of human language? We believe there is at least one such insight, and that is that contact clearly does not necessarily simplify grammatical structure, which is the impression gained for example through Pidgin and Creole studies, but that contact on the contrary tends to make things rather complicated. The competing structures in the language, the conditions under which they are allowed to coexist, and the natural ways by which competition is resolved — all these provide a challenge to the contributors to this volume, who appreciate Romani in both its unity and diversity.

Endnotes

1.   Earlier versions of the papers by Boretzky, Bubeník, Friedman, Hancock, Igla, Matras, and partly Hübschmannová & Bubeník were presented at the Second International Conference on Romani Linguistics in Amsterdam in December 1994. The contribution by Cech & Heinschink stems from a presentation at the annual meeting of the Gypsy Lore Society in Leiden in March 1995, that of Elšik and partly that of Hübschmannová & Bubeník from papers given at the Third International Conference on Romani Linguistics in Prague in December 1996.

Appendix: Dialects and varieties of Romani mentioned in the present volume

Ajia Varvara. Vlach dialect spoken in a suburb of Athens called Ajia Varvara; shows Turkish influence (see Igla 1996).
Arabadži. Balkan dialect close to Bugurdži (see Boretzky 1993).
Arli. Self-designation of a rather diverse group, mostly in former Yugoslavia. The name is derived from Turkish *yerli* 'settled'. The dialects of Prilep and Prizren (Terzi Mahalla, Yeni Mahalla) seem to be somewhat deviant dialects of Arli. The Arli dialects belong to the Balkan group. A grammatical sketch has been published by Boretzky (1996c).
Baruči(sko). See Arli. (Arli dialect of a Romani neighbourhood in Skopje, Macedonia).
Bašaldo. Self-designation of a Central dialect. See Hungarian Romani.
Bohemian. External name for a Central dialect spoken in 19th century Czechia.
British Romani. Cover term for Welsh, English, and Scottish Romani. The latter two are Para-Romani. Lexically Northern. See also English Romani.

Bugurdži. Self-designation of a Balkan dialect spoken in former Yugoslavia (see Boretzky 1993).

Burgenland Roman. A Central dialect related to the *Prekmurje* dialect of Slovenia. The Austrian variant is described in Halwachs (1996).

Burgudži. Alternative name for *Bugurdži*.

Caló. Self-designation of a Spanish-based Para-Romani. See Boretzky (1992), Bakker (1995), and Leigh (forthcoming).

Carpathian. An external label applied to a Central dialect. See Hungarian Romani.

Čurari. A Vlach dialect closely related to *Kalderaš*.

Drindari. Self-designation of a rather deviant Balkan dialect. Described in Kenrick (1969).

Džambaz(i). Self-designation of a Vlach dialect spoken in former Yugoslavia, and sometimes also called *Gurbet*.

English Romani. An English-based Para-Romani, with a Northern lexicon. In earlier stages, however, there were also English Romani dialects with the inherited grammatical system. See Hancock (1984).

Erli. See also *Arli*. Self-designation of a Balkan dialect. *Erli* is usually used for the dialect of Sofia, described by Kostov (1963) and others.

Grekuri. A Vlach dialect reported from Argentina.

Gurbet(i). Vlach dialects of former Yugoslavia, so designated to refer to recent immigrants to the region, in contrast with the 'settled' speakers of a Balkan dialect. See Boretzky (1986).

Hungarian Romani. External label of a Central dialect (but see *Romungro*). See Hancock (1990).

Kalajdži. Self-designation of a Balkan dialect spoken in Bulgaria and Turkey. See Gilliat-Smith (1935).

Kalderaš. Also *Kelderaš*. Self-designation of the Vlach dialect of the Coppersmiths. See Gjerdman & Ljungberg (1963), Boretzky (1994).

Lešaki. Self-designation of a variety of the dialect of the *Polska Roma* spoken in central Poland, and part of the Northern group of dialects. It is related to the North Russian dialect described by Wentzel (1980).

Lovari. Self-designation of a Vlach dialect spoken in Transylvania and elsewhere in central Europe. See Pobožniak (1964).

Mačvano. Vlach dialect originally spoken in Serbia, now mainly in the USA.

Náwari (also *Nuri*, *Domari*). Indic language related to Romani, spoken in the Near East. See Macalister (1909-1913).

Non-Vlach. External cover term usually used to refer to the Balkan branch.

Para-Romani. Cover term for languages which have a Romani lexicon, but a non-Romani grammatical structure. Examples are Caló and most varieties of British Romani. See Boretzky & Igla (1994b), Matras (forthcoming).

Paspatian. A Balkan dialect, called so in the literature as it has been thoroughly described by Paspati (1870). Also called *Rumeli*.

Prekmurje. See Burgenland.

Prilep. A Balkan dialect belonging to Arli, named after the town in which it is spoken.

Rabadži. See *Arabadži*.

Rhodopes. Deviant dialect of Erli spoken in the Rhodopes mountains in Bulgaria. See Igla, this volume.

Romungro. One of the self-designations of a Central dialect spoken in Hungary. The term is also occasionally used by Vlach speakers in central Europe to refer to the Northern dialect of the *Polska Roma*. See Vekerdi (1981).

Rumeli. See *Paspatian.*

Scottish Romani. An English-based Para-Romani, spoken in Scotland.

Sepeči(des). Self-designation of the Balkan dialect of the basket-weavers in Greece and Turkey. Related to Arli. See Cech & Heinschink, this volume.

Sinti. Self-designation of German Romani, a sub-branch within the Northern group of dialects. See Holzinger (1993).

Slovak Romani. External term for a Central dialect, now also spoken in the Czech Republic. See Hübschmannová et al. (1991) as well as Elšik, this volume.

Terzi Mahalla. A Balkan dialect spoken in Prizren, which can be considered a deviant variety of the Arli-type.

Vend. A Central dialect spoken in Hungary. See Vekerdi (1984).

Vlach, Vlax. External cover term for the dialects of Rumania, Transylvania, and adjoining regions, but also widespread elsewhere since the outwards migration from these areas in the nineteenth and early twentieth century. See Hancock (1995).

Welsh Romani. A sub-branch of the Northern group, now though to be extinct. See Sampson (1926).

Xoraxane. External cover term as well as collective self-designation for the dialects of Muslim Roma in the Balkans, usually heavily influenced by Turkish. It refers most frequently to dialects of the Balkan group, though occasional use of the term with reference to Vlach varieties have also been recorded.

Zargari. Self-designation of a (European, probably Balkan) dialect of Romani spoken in Iran by out-migrants. See Windfuhr (1970), cf. also Djonedi (1996).

References

Adelung, Friedrich von. 1815 [1976]. *Catherinens der Grossen Verdienste um die vergleichende Sprachenkunde.* St. Petersburg: Friedrich Drechsler. [Hamburg: Buske].

Bakker, Peter. 1995. 'Notes on the genesis of Caló and other Iberian Para-Romani varieties'. In: Yaron Matras, ed. *Romani in contact. The history, structure and sociology of a language.* Amsterdam: Benjamins. 125-150.

---- & Marcel Cortiade, eds. 1991. *In the margin of Romani. Gypsy languages in contact.* Amsterdam: Institute for General Linguistics.

---- & Yaron Matras. Forthcoming. *An indexed bibliography of Romani linguistics.*

Boretzky, Norbert. 1986. 'Zur Sprache der Gurbet von Priština (Jugoslawien)'. *Giessener Hefte für Tsiganologie* 3:1-4. 195-216.

--- 'Zum Interferenzverhalten des Romani'. *Zeitschrift für Phonologie, Sprachwissenschaft und Kommunikationsforschung* 42:3. 357-374.

---- 1992. 'Romanisch-zigeunerische Interferenzen (zum Caló)'. In: Jürgen Erfurt, Benedikt Jessing & Matthias Perl, eds. Prinzipien des Sprachwandels I. Bochum: Brockmeyer. 11-38.

---- 1993. *Bugurdži. Deskriptiver und historischer Abriß eines Romani-Dialekts.* Berlin: Harrassowitz.

--- 1994. *Romani. Grammatik des Kalderaš-Dialekts mit Texten und Glossar.* Berlin: Harrassowitz.

--- 1996a. 'The "new" infinitive in Romani'. *Journal of the Gypsy Lore Society* 6:1. 1-51.

---- 1996b. 'Entlehnte Wortstellungssyntax im Romani'. In: Norbert Boretzky, Werner Enninger & Thomas Stolz, eds. *Areale, Kontakte, Dialekte. Sprache und ihre Dynamik in mehrsprachigen Situationen*. Bochum: Brockmeyer. 95-121.

---- 1996c. 'Arli. Materialen zu einem Südbalkanischen Romani-Dialekt'. *Grazer Linguistische Studien* 46. 1-30.

---- & Birgit Igla. 1991. *Morphologische Entlehnung in den Romani-Dialekten*. Arbeitspapiere des Projektes 'Prinzipien des Sprachwandels'. Arbeitspaper Nr. 4. Essen: Fachbereich Sprach- und Literaturwissenschaften an der Universität Essen.

---- 1993. *Lautwandel und Natürlichkeit. Kontaktbedingter und endogener Lautwandel im Romani*. Arbeitspapiere des Projektes 'Prinzipien des Sprachwandels'. Arbeitspaper Nr. 15. Essen: Fachbereich Sprach- und Literaturwissenschaften an der Universität Essen.

--- 1994a. *Wörterbuch Romani-Deutsch-Englisch für den südosteuropäischen Raum*. Wiesbaden: Harrassowitz.

---- 1994b. 'Romani mixed dialects'. In: Peter Bakker & Maarten Mous, eds. *Mixed languages. 15 case studies in language intertwining*. Amsterdam: IFOTT. 35-68.

Cech, Petra & Mozes F. Heinschink. 1997. Sepečides Romani. München: Lincom.

De Goeje, M. J. 1903. *Mémoires sur les migrations des Tsiganes à travers l'Asie*. Leiden: E. J. Brill.

Demeter, R. S. 1990. *Cygansko-russkiy i russko-cyganskiy slovar'*. Moskva: Russkiy yazyk.

Djonedi, Fereydun. 1996. 'Romano-Glossar. Gesammelt von Schir-Ali Tehrani-zade'. *Grazer Linguistische Studien* 46. 31-59.

Finck, Franz Nikolaus. 1907. 'Die Grundzüge des Armenisch-Zigeunerischen Sprachbaus'. *Journal of the Gypsy Lore Society* 1. 34-60.

Friedman, Victor A. 1985. 'Balkan Romani modality and other Balkan languages'. *Folia Slavica* 7:3. 381-389.

--- 1995. 'Romani standardization and status in the Republic of Macedonia'. In: Yaron Matras, ed. *Romani in contact. The history, structure and sociology of a language*. Amsterdam: Benjamins. 177-188.

---- & Robert Dankoff. 1991. The earliest text in Balkan (Rumelian) Romani: A passage from Evliya Çelebi's Seyāḥat-nāme'. *Journal of the Gypsy Lore Society* 1:1. 1-20.

Gilliat-Smith, Bernard J. 1915-1916. 'Report on the Gypsy tribes of north-east Bulgaria'. *Journal of the Gypsy Lore Society* 9. 1-55; 65-108.

---- 1935. 'The dialect of the Moslem Kalajdžis (Tinners) of the Tatar Pazardžik district'. *Journal of the Gypsy Lore Society* 14. 25-43.

Gjerdman, Olof & Erik Ljungberg. 1963. *The language of the Swedish copper-smith Gipsy Johan Dimitri Taikon*. Uppsala: Lundequist.

Grellman, Heinrich M. G. 1787. *Historischer Versuch über die Zigeuner, betreffend die Lebensart und Verfassung, Sitten und Schicksale dieses Volkes seit seiner Erscheinung in Europa und dessen Ursprung*. Göttingen: Dietrich.

Haarmann, Harald. 1989. [Introduction to reprint of Rüdiger 1782]. Hamburg: Buske. vii-xxvii.

Halwachs, Dieter W. 1996. *Morphologie des Roman. Basisgrammatik der Romani-Variante der Burgenland-Roma*. Oberwart: Verein Roma.

Hancock, Ian F. 1976. 'Patterns of English lexical adoption in an American variety of Romanes'. *Orbis* 25:1. 83-104.

--- 1984. Romani and Angloromani. In: Peter Trudgill, ed. *Languages in the British Isles*. Cambridge: Cambridge University Press. 367-383.

---- 1987. *The pariah syndrome. An account of Gypsy slavery and persecution.* Ann Arbor: Karoma.

--- 1988. 'The development of Romani linguistics'. In: Mohammad Ali Jazayery & Werner Winter, eds. *Languages and cultures. Studies in honor of Edgar C. Polomé*. Berlin: Mouton de Gruyter. 183-223.

---- 1990. *A grammar of the Hungarian-Slovak (Carpathian, Bashaldo, Rumungro) Romani language*. Manchaca: International Romani Union.

---- 1995a. *A handbook of Vlax Romani*. Columbus: Slavica.

---- 1995b. 'On the migration and affiliation of the Ḍōmba: Iranian words in Rom, Lom and Dom Gypsy. In: Yaron Matras, ed. *Romani in contact. The history, structure and sociology of a language*. Amsterdam: Benjamins. 25-51.

Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Holzinger, Daniel. 1993. *Das Rómanes: Grammatik und Diskursanalyse der Sprache der Sinte*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Hübschmannová, Milena, Hana Šebková & Ana Žigová. 1991. *Romsko-český a Česko-romský kapesní slovník*. Praha: Státní pedagogické nakladatelství.

Igla, Birgit. 1996. *Das Romani von Ajia Varvara: Deskriptive und historisch-vergleichende Darstellung eines Zigeunerdialekts*. Wiesbaden: Harrassowitz.

Kenrick, Donald. 1969. *Morphology and lexicon of the Romany dialect of Kotel (Bulgaria)*. Ph.D. dissertation, School of Oriental and African Studies, London.

Kluge, Friedrich. 1901 [1987]. *Rotwelsch. Quellen und Wörterbuch der Gaunersprache und der verwandten Geheimsprachen*. Straßburg: Karl Trübner. [Berlin: de Gruyter].

Kostov, Kiril. 1963. *Grammatik der Zigeunersprache Bulgariens: Phonetik und Morphologie*. Ph.D. dissertation, Humboldt University, Berlin.

Leigh, Kate. Forthcoming. 'Romani elements in present-day Caló'. In: Yaron Matras, ed. *The Romani element in non-standard speech*. Wiesbaden: Harrassowitz.

Lucassen, Leo. 1996. *Die Zigeuner. Die Geschichte eines polizeilichen Ordnungsbegriffes in Deutschland 1700-1945*. Köln: Böhlau.

Macalister, R. A. Stewart. 1909-1913. 'A grammar and vocabulary of the language of the Nawar of Zutt, the nomad smiths of Palestine. *Journal of the Gypsy Lore Society* 3. 120-126, 298-317; 5. 289-305; 6. 161-240.

Matras, Yaron. 1994a. *Untersuchungen zu Grammatik und Diskurs des Romanes: Dialekt der Kelderaša/Lovara*. Wiesbaden: Harrassowitz.

---- 1994b. 'Structural Balkanisms in Romani'. In: Norbert Reiter, Uwe Hinrichs & Jeřina van Leeuwen-Turnocová, eds. *Sprachlicher Standard und Substandard in Südosteuropa und Osteuropa*. Berlin/Wiesbaden: Harrassowitz. 195-210.

---- ed. 1995a. *Romani in contact. The history, structure and sociology of a language*. Amsterdam: John Benjamins.

---- 1995b. 'Verb evidentials and their discourse function in Vlach Romani nar-
ratives'. In: Yaron Matras, ed. *Romani in contact. The history, structure and
sociology of a language.* Amsterdam: John Benjamins. 95-123.
---- 1995c. 'Connective word order in Romani'. *Sprachtypologie und Univer-
salienforschung* 48:1-2. 189-203.
---- 1996. 'Prozedurale Fusion: Grammatische Interferenzschichten im
Romanes'. *Sprachtypologie und Universalienforschung* 49:1. 60-78.
---- In press. 'Deixis and deictic oppositions in discourse: Evidence from
Romani'. *Journal of Pragmatics.*
---- ed. Forthcoming. *The Romani element in non-standard speech.* Wiesbaden:
Harrassowitz.
Miklosich, Franz. 1872-1880. *Über die Mundarten und Wanderungen der
Zigeuner Europas.* I-XII.Wien: Karl Gerold's son.
Moravcsik. Edith A. 'Language contact'. In: Joseph H. Greenberg, Charles A.
Ferguson, & Edith A. Moravcsik, eds. *Universals of human language* Vol. I.
Stanford: Stanford University Press.
Okely, Judith. 1983. *The Traveller Gypsies.* Cambridge: Cambridge University
Press.
Paspati, Alexandre G. 1870 [1973]. Études sur les Tchinghianés ou Bohémiens de
l'Empire Ottoman. Constantinople: Koroméla. [Osnabrück: Biblio].
Payne, John R. 1995. 'Inflecting postpositions in Indic and Kashmiri'. In: Frans
Plank, ed. *Double case. Agreement by Suffixaufnahme.* New York: Oxford
University Press. 283-298.
Piasere, Leonardo. 1994. *Il piu antico testo Italiano in Romanes (1646): una
risoperta e une letture etnostorica.* Universita' degli Studi di Verona. Facolta'
di Lettere e Filosofia, Istituto di Psicologia, Report 56.
Plank, Frans. 1995. '(Re-)Introducing Suffixaufnahme'. In: Frans Plank, ed.
Double case. Agreement by Suffixaufnahme. New York: Oxford University
Press. 3-110.
Pobożniak, Tadeusz. 1964. *Grammar of the Lovari dialect.* Kraków: Państwowe
wydawnictwo naukowe.
Pott, Augustus F. 1844-1845 [1964]. *Die Zigeuner in Europa und Asien.
Ethnographisch-linguistische Untersuchung vornehmlich ihrer Herkunft und
Sprache.* Halle: Heynemann. [Leipzig: Edition Leipzig].
Ruch, Martin. 1986. *Zur Wissenschaftsgeschichte der deutschsprachigen
"Zigeunerforschung" von den Anfängen bis 1900.* Ph.D. dissertation, Uni-
versity of Freiburg.
Rüdiger, Johann C.C. 1782 [1990]. *Von der Sprache und Herkunft der Zigeuner
aus Indien. Nachdruck aus: Neuester Zuwachs der teutschen, fremden und
allgemeinen Sprachkunde in eigenen Aufsätzen, 1. Stück.* Hamburg: Buske.
Sampson, John. 1907. 'Gypsy language and origin'. *Journal of the Gypsy Lore
Society* 1. 4-22.
--- 1926 [1968] *The dialect of the Gypsies of Wales.* Oxford: Clarendon.
Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creoliza-
tion and genetic linguistics.* Berkeley: University of California Press.
Turner, Ralph. 1926. 'The position of Romani in Indo-Aryan'. *Journal of the
Gypsy Lore Society* 5:4. 145-194.
Van der Auwera, Johan (with Dónall P. Ó Baoill). ed. In press. *Adverbial con-
structions in the languages of Europe.* Berlin: Mouton de Gruyter.
Vekerdi, Jószef. 1981. *A Magyar Cigány nyelvjárás nyelvtana.* Pécs: Janus
Pannonius University.

---- 1984. 'The Vend Gypsy dialect in Hungary'. *Acta Linguistica Academiae Hungaricae* 24. 381-389.

Wentzel, Tatjana. 1980. *Die Zigeunersprache (nordrussischer Dialekt).* Leipzig: Enzyklopädie.

Wexler, Paul. Forthcoming. 'Could there be a Rotwelsch origin of Romani lexicon?' In: Yaron Matras, ed. *The Romani element in non-standard speech.* Wiesbaden: Harrassowitz.

Willems, Wim. 1995. *Op zoek naar de ware zigeuner. Zigeuners als studieobject tijdens de verlichting, de romantiek en het nazisme.* Utrecht: Jan van Arkel.

Windfuhr, G.L. 1970. European Gypsies in Iran. A first report'. *Anthropological Linguistics* 12. 271-292.