Book Reviews


Reviewed by Victor A. Friedman

Although not all Gypsies speak Romani, all native speakers of Romani as their first community language are Gypsies (in a restricted, ethnic sense of the term). For some groups of Gypsies who do not speak Romani, the evidence is clear that Romani was the primary community language at one time and that the community has subsequently undergone language shift. The term 'Gypsy', however, is sometimes also used to refer to peripatetic social groups who may or may not speak or have ever spoken Romani. In the case of some West European groups to whom the term 'Gypsy' is applied, it is certain that Romani is not and was not used as a community language. Similarly, the so-called 'Gypsy languages' of India, described, for example, by G. A. Grierson in Volume XI of the Linguistic Survey of India (1922, Calcutta, Office of the Superintendent of Government Printing), are not forms of Romani but Indic languages of India that have been separated from Romani for at least as long as all the other languages of India (more than a millennium) and whose speakers occupy socio-economic niches similar to those of some Gypsies in Europe.

In some recent West European historical and anthropological literature, a confusion has arisen between the use of 'Gypsy' as an ethnic label to refer to Romani-speaking or formerly Romani-speaking groups (who may or may not be peripatetic) and the second usage, which is more social than ethnic and does not have a firmly definable linguistic basis. Members of some groups of Gypsies in this second sense have secret lexicons that are not independent languages. Some historians and anthropologists have become so confused by these two usages of the term 'Gypsy' as to assert that Romani is merely a lexicon, not a language, and that there was no migration of speakers of Early Romani to Europe from India but rather that peripatetic peoples passed Indic lexical items from one group to another along trade routes.

Yaron Matras’s Romani: A Linguistic Introduction (henceforth RLI) is a milestone in the past two centuries of linguistic scholarship on Romani that has demonstrated conclusively that Romani is indeed a language, that it is

Victor A. Friedman is member of the editorial board of Romani Studies and Professor of Slavic Languages and Linguistics at the Slavic Department, University of Chicago, 1130 East 59th Street, Chicago, IL 60637, USA. Email vfriedman@midway.uchicago.edu

Romani Studies 5, Vol. 13, No. 2 (2003), 165–173 ISSN 1538–0478
Chapter 2 gives a very brief overview of the dialectal classification of Romani currently accepted by most linguists working on Romani. The four living branches are Balkan, Vlax, Central, and Northern, each of which is divided into two subgroups—southern and northern for the first three and eastern and western for the last. Matras also mentions British/Welsh and Iberian Romani, each considered as forming independent branches and surviving only as special lexicons embedded in territorial languages. Matras also treats the Abruzzian and Hrvati/Dolenjški dialects as taxonomically distinct, although it appears that they are basically southern Balkan dialects that separated early from the main group and subsequently underwent significant changes under the influence of other centers of innovation and territorial languages. I shall return to this issue in my discussion of Chapter 9, to which Matras refers the reader in his mention of these dialects. Chapter 2 concludes with a geographic and a schematic map of the main dialects that serve as the bases of discussion in RLI and a brief mention of Para-Romani, the Romani-based lexicons that function as special variants of territorial languages and that are discussed in Chapter 10. Owing to the dialectal diversity of Romani and the fact that it is spoken all over Europe (and, as a result of subsequent migration, on most other continents as well), particularly by some populations that are or were peripatetic, Romani is described as a non-territorial language, although specific dialects are associated with geographic locations. Any linguistic study of Romani must take into account its dialectal diversity, since there is no dominant standard that can serve as a point of reference. Chapter 2 thus serves as the basic frame for the rest of the discussion. Since RLI is indeed an introduction and not an exhaustive compendium, and, moreover, since the documentation of Romani dialects and their nomenclature is the subject of ongoing research (most notably the Romani Morpho-Syntactic Database Project at the University of Manchester directed by Yaron Matras and the Romani Lexical Project, directed jointly by Dieter Halwachs at the University of Graz, Peter Bakker at the University of Aarhus, and Yaron Matras at the University of Manchester), RLI does not attempt to provide a complete account of all Romani dialects but rather an inclusive overview. Thus, for example, Figure 2.1, Location of the principal dialects of Romani, leaves Albania completely blank not because Romani is not spoken there, nor because we lack documentation for some of the local dialects, but because the principal dialectal types needed to form the basis of discussion in RLI are represented by other well-documented dialects spoken in adjacent countries.

Indic, and that the only reasonable hypothesis for its presence in Europe is that a group or some groups of speakers migrated thither from India. But Matras’s book is not really concerned with refuting the claims of a few misguided Western scholars unacquainted with the field of linguistics, although the material presented in it does so. Rather, this volume gives both a state-of-the-art survey of current knowledge relating to the Romani language and introduces some of the author’s original ideas concerning various aspects of Romani linguistics. RLI is not only the first such work to be published in English but is also the best of its kind in the field. It is essential reading for anyone concerned with Roms or the Romani language, but it also has much to contribute to general linguistics and can be used with profit by scholars in other areas of the humanities and the social sciences.

The book is divided into eleven chapters: 1. Introduction (pp. 1–4), 2. Romani dialects: a brief overview (5–13), 3. Historical and linguistic origins (14–48), 4. Descriptive phonology (49–71), 5. Nominal forms and categories, 6. Verb morphology (71–64), 7. Syntactic typology (65–90), 8. Grammatical borrowing (91–213), 9. Dialect classification (214–37), 10. Romani sociolinguistics (338–50), and 11. Language planning and codification (251–59), followed by a bibliography of over 450 references (260–78) and indices of dialects (279–81), names (282–85), and subjects (286–91). Chapters 4–7 constitute the core of the synchronic description of the structure of Romani combined with relevant diachronic explanations. These chapters are aimed specifically at linguists and require at least a basic knowledge of linguistics. The theoretical orientation is functional–structural, and the material and analyses are accessible to any linguist, clearly and coherently presented, and contain a number of new and convincing explanations. The first three chapters and the last four are concerned with those aspects of Romani linguistics that are of interest to both linguists and scholars in other fields: They can be assigned independently of Chapters 4–7 in any course for which Roms in particular or Gypsies in either sense constitute the focus of study or a relevant subject.

Chapter 1 introduces Romani and gives a brief overview of the history of its study, mentioning all the earliest accounts up to the eighteenth century and then surveying the most important figures and events up to the present. The author concludes this chapter by addressing the ethical responsibility of researchers to the community that is the subject of research and assessing in a realistic fashion the contributions that Romani linguistics can make to the Romani-speaking community.
words relating to areas of contact with surrounding populations are more likely candidates for loan-replacements while those dealing with more intimate spheres show greater stability and resistance to replacement. He cites Boretzky’s (1992) conclusion that Romani’s relatively small common inherited vocabulary flags resistance to assimilation. Matras gives a fairly detailed account of Romani historical phonology, a brief survey of the most important features of historical morphology, and a discussion of the relationship of Romani to Domari and Lomavren. He supports Turner’s (1926) arguments for the origin of all three languages in Central India. Matras also notes that Turner failed to take into account the Romani preservation of a distinct reflex for retroflex d, and, citing the reflexes of Old Indic bhīraṇa- (‘brother’) (Romani plhāl, Lomavren plhāl, Domari bar) along with other phonological, morphological, and lexical evidence, allows for the possibility that the three languages became distinct as a result of their emergence as non-territorial languages in Central India, with subsequent shared innovations resulting from areal contact as the groups followed similar paths of migration through northwestern India while perhaps separating completely upon reaching Iranian-speaking territory.

The descriptive linguistic core of RLI, Chapters 4 through 7, discuss phonology, the nominal system, the verbal system, and syntactic typology. The opening treatment of palatals in Chapter 4 is a bit confusing, since /c/ is described as behaving differently from other stops (emphasis mine), but later in the paragraph Matras points out the need to distinguish palatal stops, palatalized consonants, and affricates. It would have been better to formulate this description in terms of four primary positions for obstruents (labial, dental, palatal, velar), three of which always have stops. Otherwise the chapter gives a thorough survey of Romani consonantism, vocalism, and prosody as well as the main morphophonemic alternations: prothesis/truncation, jotation, and the alternation of /s/ with /sh/, /f/, or zero. The treatment of this last alternation gives a much needed synthesis of a complex phenomenon and is based in large part on Matras’s own research. For this alternation, which actually involves two phenomena—an older one in intervocalic position and a younger one in final position (the former restricted to certain grammatical morphemes, the latter either grammatical or phonological depending on the dialect)—Matras establishes a hierarchy of occurrence and argues that the older phenomenon can be traced back to variation in Proto-Romani and Early Romani beginning with the second-person singular present-tense marker (-esa) and spreading by analogy first.

Chapter 3 surveys Romani historical linguistics and its contribution to the reconstruction of Romani history. Matras’s sober and careful account makes it clear that while historical linguistics establishes the origins of Romani in India beyond the shadow of a doubt, and moreover gives us some indications of the time and route of migration, it cannot tell us the reason for that migration, nor can it establish the social position of the speakers whose language became Romani. As Matras points out, the ethnonym Rom (Rom) is certainly cognate with the attested caste name Dom, which refers to various service-providing populations in India and is also cognate with ethnonyms of Indic-speaking or formerly Indic-speaking commercial nomads (some of whom use other names) who live all over central and western Asia and adjacent parts of Africa from Tajikistan and the Caucasus to Egypt and Sudan. Having discussed all the theories of Romani’s social origin, Matras concludes that what we know for certain is that mobile populations of Indic origin specializing in certain trades and retaining distinct ethnic and linguistic identities were present in the Near East before the Rom arrived in Byzantium ‘sometime around the eleventh century or earlier’ (18). Worth citing here is evidence from Tzitzilis (2001), who argues on the basis of the phonology and morphology of Middle Greek loan-words that Romani was being spoken in Byzantium by the tenth century. Note particularly that Tzitzilis (2001: 327) observes that the Greek source of Romani kurkā is not the usually cited kyrīaki but a different expression for ‘Sunday’ kyrīkōn (imur) (pace p. 205).

Matras defines Proto-Romani as the Indic dialect that became a distinct language in India and Early Romani as the language that was spoken in contact with Greek in Byzantium prior to the dispersal of Romani-speaking populations throughout Europe beginning in the thirteenth or fourteenth centuries. These are distinct from Common Romani, which is the term he uses for those Early Romani elements shared by attested Romani dialects, as opposed to variations that may have been present in Early Romani or Proto-Romani. In his discussion of the Early Romani inherited lexicon, Matras identifies about 1,000 lexical roots, of which between 200 and 250 are from Greek, around 70 from Iranian, about 40 from Armenian, leaving about 650 to 700 roots of Indic origin plus a few items from Georgian, Ossetian, and, perhaps, Burushaski. Matras discusses some of the most significant examples as well as their semantic domains (kinship terms, body parts, etc.) and the question of what this lexicon can tell us about Romani culture. After a short but careful survey of conflicting hypotheses, Matras concludes that
sets in terms of two oppositions: remote–non-remote and perfective–non-
perfective. The remoteness marker is -as (sine in the Arli of Kosovo and
Macedonia) and the perfective stem uses the participial base and person
markers of pronominal origin (see below). The four paradigms thus de-
fined are the present (–R –P), imperfect (+R –P), preterite or aorist (–R +P),
and pluperfect (+R +P). The pluperfect can also function as a counterfac-
tial. Matras has a third non-remote category, intentional, which he labels as
marking the subjunctive, for example, long kerava 'to' isg present vs short
kerav isg subjunctive. While this may have been the case in Early Romani,
such a description is problematic in the modern dialects that do not use
the long form as the future. Although it is common for the short present to
occur after modal markers such as the future marker ka and the modal sub-
ordinator te, it is also the case in dialects such as Skopje Arli that the long
form occurs after te when introducing the protasis of a conditional period
as well as after bi marking the apodosis, for instance, Te tigareu man, me bi
avava 'If they were to invite me I would come'. Short forms can also occur in
simple statements—for example, En rifatos pendzarev, e čhaja da pendzarev
'I know Rifat and I know his daughter'. Consider also the following radio an-
nouncement, which has long and short indicative: O Ajnur thaj o Džemo
tar-i Svedska baktare e pranden e Ramjiske thaj e Mirdakate, a e Sajetske
thaj e Sadijake baktarena o bijav Ajnur and Džemo from Sweden congrat-
ulate Rami and Mirsada on their marriage, and they congratulate Safet and
Sadija on their wedding.' Thus the characterization in Table 6.12 (p. 156)
necessitates further explication. Matras's explanation for the origin of the simple
preterite in an ergative-like construction going back to a past participle plus
a possessive or relative linking particle 'jo plus a pronoun is original, ingen-
ious, and convincing. It gives the first truly satisfactory account for both jota-
tion and person-marking in the preterite.

Chapter 7 illustrates the most important syntactic features of Romani
using 96 example sentences, all taken from unelicited data sets. Matras's ac-
count details the syntax of noun phrases, constituent order in verb phrases,
possession, complex clauses, and negation. The discussion of Romani syn-
tax brings together much varied material, and this is perhaps the most com-
plex and difficult area of Romani to survey. Thus, for example, Macedonian
deka is also used as a factive complementizer in Arli, and Macedonian could
have been the intermediary for Greek oti, which also occurs in Macedonian
(cf. p. 179), Balkan Turkish iger, 'if', is used alongside South Slavic ako in
Macedonia (cf. p. 187), etc. Matras makes a useful distinction between stable,
to the first-person plural present, then the instrumental singular, the cop-
ula, (third-person singular present before other forms), and interrogatives,
in various dialects which tend to cluster in Central Europe. Loss of final /a/
in Greek-derived nouns, the third-person singular past marker (-as), the re-
flexive marker (pres), masculine accusative (independent oblique) singular
(-es), adverbial -es, the short present (second-person singular, first-person
plural), and inherited lexical items represents different processes in different
dialects and also follows a hierarchy of occurrence. The most conservative
are those at the periphery.

Chapter 5 treats the inflection and derivation of substantives, adjectives,
and deictics (pronouns and articles). In discussing case markers, Matras
231–2, 234, 236), Layer I markers are 'bits of material inherited from M[iddle]
I/nd0-][Aryan]/O[ld][I/nd0-][Aryan];' Layer II markers are 'attached to
the base indirectly, through the mediation of a Layer 1 element and/or invariant
for all nouns and the same for all numbers.' Layer III markers are mediated
by Layer II, and, when they are not, are still distinguishable by their lack of mor-
phophonemic variation, ability to be polysyllabic, and, usually, a 'fairly trans-
parent connection with an independent word.' Masica needs this system owing
to the fact that the Indic languages of India are all postpositional and
therefore the interface between the three layers requires delicate distinctions.
Matra's use of Masica's terminology is justified, for the most part, on his-
torical grounds, since most of the Romani substantival inflectional system
derives from the innovations of the Middle Indic period that led to the for-
mation of the first two layers. In Romani, however, the material correspon-
ding to Masica's Layer III consists of prepositions which are always identifiable
as independent words and frequently do not trigger the occurrence of Layer
II markers (except in pronouns or in some dialects in contact with languages
having complex case inflection). Moreover, the Romani vocative markers, as
Matras notes (p. 80), connect directly to the base but are of late formation
and do not belong to Layer I (nor to Layer II). Matras's treatment of the in-
dependently occurring oblique stem (frequently labeled 'accusative') as 'en-
coding the non-agentive referent that is high on the topicality scale' (p. 86)
captures the fact that this form is not limited to animate or definite direct ob-
jects but also occurs in other functions. Matras's account of the development
of Roman's distincitively complex set of deictics is original and convincing.

Chapter 6 gives a thorough account of Romani inflectional and der-
ivational verb morphology. Matras analyzes the four basic paradigmatic
Common Romani Balkanisms and those Balkan contact features that are specific to the dialects of Southeastern Europe. In the concision, he identifies three Early Romani innovations as typical of Romani syntax: (1) relative clauses with resumptive pronouns and interrogatives as relativizers; (2) a non-factive–factive distinction in subordinators (native te for the former in almost all dialects, native kaj [‘where’] or non-native forms for the latter); and (3) predominant VO word order with OV for focus, VS for connective-narrative, and SV for contrastive thematic orders.

Chapter 8 discusses Romani in the context of the field of contact linguistics. The origins of Romani are quite distinct from what is normally understood by the term contact language–namely, a language coming into existence as the result of contact between two or more speech communities. Rather, as Matras observes, Romani is a language ‘permanently in contact’ (p. 191) with other languages. There are no monolingual speakers aside from very young children, and until the twentieth century there were no organized efforts at language preservation. While there are other speech communities in the world where bi- or multilingualism is the universal, Romani is, perhaps, unique in its history of non-territoriality combined with what must be assumed to have been centuries of universal multilingualism. Matras discusses Romani evidence in the context of various theoretical attempts to predict the relative borrowability of grammatical elements. Using Romani data he establishes an implicational hierarchy of the likelihood of borrowing in conjunctions (from least to most frequently borrowed and/or replaced): ‘and’ < ‘or’ < ‘but’. Other areas of borrowing are discourse markers, subordinating conjunctions of reason, cause, factivity, and time, various expressions of modality, adjectival gradation, bound indefinite markers, focal quantifiers, and bound morphemes used in nominal derivation. In nominal inflection, Matras mentions the borrowing of the nominative plural markers from Romanian (-urj/-uri), Greek (-ides), and Greek–Slavic -i. He also cites the use of -e in the Balkans as being of South Slavic origin, but it is unclear why this could not be an extension of the plural marker for native masculine nouns in -o, albeit possibly with South Slavic reinforcement. We can also mention here the influence of Slavic inflectional patterns on adjectival agreement, which Matras notes on p. 96. Borrowed prepositions generally express abstract rather than concrete spatial relations. Matras also notes a variety of phonological influences from contact languages (mainly stress and vowels) as well as syntactic convergences. Among those elements which are not borrowed and thus represent significant constraints are the follow-
References


a taxonomy assigning a systemic position, if only for heuristic purposes. Be that as it may, Matras's original contributions and coherent synthesis constitute the best and most up-to-date general account of Romani dialect classification currently available.

Chapter 10, on Romani sociolinguistics, discusses the problem of estimating the actual number of Romani speakers, issues of bilingualism, code-switching and code-mixing, Para-Romani, and the Romani element in other languages. Matras makes the important point that the tendency for Romani words to be recruited into slang, especially in the expression of subjects that are pejorized or taboo in mainstream society, is an expression of the high prestige of Romani for anti-establishment, marginalized communities. As slang itself is a challenge to established norms, the Romani element in slang indicates its significance for challenges to the social order.

On the other hand, the final chapter of RLI treats Romani as a participant in the established social order, namely, through codification and language planning. Matras covers the main issues of this question and shows that the Romani standardization process is decentralized and pluralist. It is worth noting that the situation in Macedonia, where efforts at creating a Romani standard go back to the early 1970s, current trends in language planning belie traditional linear or even circular models of the relevant processes and require more complex analyses. Matras concludes with a discussion of the status of Romani and notes the rise in its recognition in academic research contexts and the importance of Romani web sites in increasing Romani's presence in the public domain. To this we can add that the achievement of Matras's book itself is a significant event in the advancement of the status of Romani.

The book will be of tremendous use to a broad range of scholars—not only linguists but also social scientists and other students in various humanistic disciplines. There is nothing like it available in any language, and it fills a pressing need. RLI's theoretical frameworks are sound, the explanations are well argued, and the coverage is thorough. Matras discusses every major topic of concern to Romani and of relevance to Romani's relation to a broad range of theoretical linguistic issues. This book is not merely a collection of already established knowledge but an original contribution to the advancement of both Romani scholarship and a variety of theoretical concerns within linguistics.