ELSEVIER

Book review

Yaron Matras, Romani: A Linguistic Introduction, Cambridge University Press, Cambridge, 2002; (pp. viii–xiii and 1–291, hardback, ISBN 0 521 63165 3).

1. Introduction: defining Romani

Romani is the only Indic language spoken exclusively outside the Indian subcontinent, mainly in Europe, since the Middle Ages. The term *Romani* is an endo-linguonym which native speakers – who generally call themselves *Roma*, while non-Roma often call them Gypsies – use to designate their language. The word *řomani* is the feminine singular adjective derived from the noun *řom*. As a linguonym it is used in the feminine gender because the noun *čhib* 'language' which it modifies, is feminine.

Modern linguistic literature, conforming to the native-language self-definition of communities speaking this language, generally use the term *Romani* to refer to this Balkanized Indic language, or some variety thereof (cf. Matras, 1994a; Friedman, 1985). Romani is the language of a non-territorial, transnational ethnic group with no political state. The number of speakers is estimated to be somewhere between 3.5 and 10 million.

2. Romani linguistic research[1]: traditions and paradigms

2.1. "Old generation Romani linguistics"

The earliest language data on Romani dates from the 16th century, and a number of sources, typically wordlists or short sentences, are known from the 16th and 17th centuries. The beginning of linguistic research on Romani may be placed in the second half of the 18th century (Matras, 1999a). At that time linguists dealt with Romani primarily as an Indic language. Called "old generation Romani linguistics" by Yaron Matras, this era

---

[1] Romani-related linguistic research between 1900 and 2003 is a marvellously represented in a bibliography of more than 2500 items edited by Bakker and Matras (2004). The bibliography extends to two additional Indic languages related to Romani: Domari, spoken by peripatetic communities in the Middle East, and Lomavren, a special lexicon used by the Lom living in Armenia and eastern Turkey.

achieved two discoveries still valid today: the formulation of hypotheses on the Indic origin of Romani, and the distinction between Romani as an independent, inflected language and the argot or traveller's jargon containing Romani elements.

2.2. Post-war, modern research on Romani linguistics

Numerous descriptive works were produced in the second half of the 20th century. The last few decades of the century brought a new revival of Romani linguistic research, with a broadening of the research agenda, critical analysis and improvement of earlier research findings, international projects, and wide-scale cooperation-based projects in dialectology, morphosyntax and lexicography. The International Conference on Romani Linguistics has been held regularly since 1993, publishing numerous edited volumes of its presentations.

As demonstrated by the number of works published in the last 15 years as well as by the structure of the book reviewed here, the most significant advances have been primarily in typology, dialectology (cf. Matras et al., 1997) contact linguistics (Bakker and Cortiade, 1991; Matras, 1995a, 1998a) and studies of the structure of Romani in terms of general linguistic theories (Elšik and Matras, 2000). Certain aspects of language use, particularly issues of language planning and language policy have likewise been a focus of interest in recent decades, partly as a result of Romani political and cultural movements.

3. Some general remarks

This 11-chapter book provides considerably more than the modest subtitle implies: more than a mere introduction to Romani linguistics, it is a systematic, comprehensive survey in a number of senses. It is complex in that it simultaneously undertakes the examination of Romani as a well-defined entity characterized by structural unity as well as a presentation of the high degree of internal diversity which characterizes Romani as a non-territorial minority language. The volume examines the linguistic facts of unity and variability from both synchronic and historical perspectives.

The book is also comprehensive with regard to its use of a huge linguistic database, ranging from the scattered data of older works in descriptive and comparative linguistics to the latest results of recent data collection projects. The language data spans a broad spectrum not only chronologically but also geographically (from the southern Balkans to Finland). The book is not limited to the study of a selected dialect or dialect group of Romani, but analyzes and compares data from a number of different Romani dialects. Finally, the work is comprehensive yet again in that it provides a good representation of the progress that has taken place in the above-mentioned areas of Romani-related research in the last 20 years. As a result, Matras' work will serve as an invaluable reference book for anyone interested in Romani, regardless of which dialect or linguistic level they are dealing with.

After a short introductory chapter, Chapter 2 surveys the dialects discussed in the book, as well as their groupings and related terminology. Chapter 3 focuses on historical questions. It gives a detailed analysis of the phonological system and changes in the morphology and lexicon, from which it reconstructs the features of Proto-Romani and

Early Romani. The next four chapters examine the current structural issues of Romani, by discussing in detail the phonology, morphology (Chapter 5 Nominal forms and categories, Chapter 6 Verb morphology) and syntactic typology. Chapter 8 is devoted to the study of the main patterns of grammatical borrowing. Chapter 9 outlines a system of dialect classification from a new perspective. The last two chapters deal with issues related to language use. Chapter 10, Romani Sociolinguistics, analyzes the sociolinguistic situation of Romani-speaking communities, Para-Romani varieties, and the influence of Romani on other languages and varieties; Chapter 11 examines issues of language policy and language planning.

The author is a professor of linguistics at Manchester University. In addition to descriptive linguistic works he has published numerous studies on issues of dialectology, contact linguistics, language typology, general linguistics and pragmatics. His language focus is on Romani, dialects of German, Domari, Kurdish and other languages of the Near East, and mixed and secretive languages. He has served as editor on a number of volumes presenting the latest findings in Romani linguistic research. From 2000 he serves as the editor of the journal Romani Studies. He has taken significant part in several international research projects on Romani linguistics, including the planning and coordination of the Morphosyntactic Typology of Romani Dialects and the Romani Lexical Database Project (co-ordinated by Yaron Matras, Peter Bakker and Dieter Halwachs).

4. Discussion of the various chapters

After a brief introductory chapter on terminological and research history aspects, Chapter 2 confronts the reader with the dialectal diversity of Romani. The chapter uses the classification model later to be outlined in Chapter 9 to group hitherto documented dialects of Romani. In recent years literature on Romani dialectology and contact linguistics has reached a consensus of sorts in the classification (cf. Bakker and Matras, 1997) of internal variability of Romani. In this context four major dialect groups are distinguished: the Balkan, Vlax, Central and Northern dialect groups.

The Balkan dialects are characterized by continuous, considerable Greek influence over a longer period than Romani varieties in the other three groups, whose speech communities left the Balkans earlier, presumably in the 14th and 15th centuries at the end of the Early Romani period. The Balkan dialects exhibit a strong Turkish influence in addition to the Greek. Within the Balkan branch a more conservative southern group is identified. This latter includes the Arli dialect, the largest in the region in terms of number of speakers and geographical spread, spoken by Romani communities living mainly in Greece, Macedonia and Albania. The Southern Balkan dialects also include the recently documented Sepečides variety, spoken in northern Greece and Turkey; Erli, spoken in Sofia; and Crimean Romani, spoken today mainly in Georgia. A group labelled as the Drindari-Kalajdži-Bugurdži group (Boretzky, 2000), thought to emerge in northeast Bulgaria, constitutes the other subgroup of the Balkan dialects.

The Vlax dialect group is one of the most important with regard to number of speakers, geographical spread and level of documentation. A common feature of dialects in this group is a strong Romanian influence on the lexicon, phonology and loan morphology, as

well as numerous shared internal innovations. The migration of groups speaking Vlax Romani varieties is connected with the abolition of serfdom in Romania in the late 19th century. A southeastern subgroup of the Southern Vlax dialects is comprised of Vlax Romani varieties spoken in southern Greece, while the southwestern subgroup consists of the so-called Gurbet-type Romani varieties, documented to be spoken mainly in Serbia, Bosnia, Albania and Kosovo, in addition to Xoroxano, a vernacular variety of Muslim Roma who migrated to Italy. The two best-known dialects in the Northern Vlax group are Kelderaš and Lovari. The majority of Romani native speakers in Hungary speak the Lovari variety. Available grammatical descriptions on these two dialects have mostly studied Romani communities that migrated to Western Europe (e.g., Matras, 1994b) or America. The less-studied varieties of Čurari and Cerhari (spoken in northeastern Hungary) may also be classified in the Northern Vlax subgroup, although Cerhari is a transitional variety, bearing a number of features shared by Central dialects.

The Northern subgroup of the Central dialect group includes West Slovak and East Slovak Romani as well as varieties spoken in southern Poland and the Transcarpathian Ukraine. The Northern Central dialects retain a layer of Hungarian influence. The dialects classified into the other, Southern Central subgroup are often referred as the-*ahi* dialects, as a result of their uniquely characteristic imperfect/pluperfect suffix. The varieties in this subgroup with a strong Hungarian influence include Romungro, spoken mainly in Romani communities in southern Slovakia (Elšik et al., 1999) as well as a small number of communities in Hungary, and the so-called Vend Romani varieties: Vend Romani variety spoken in Hungary, the Prekmurje variety of northern Slovenia, and Roman the Romani dialect documented in Burgenland (Halwasch, 1998), Austria.

The label Northern Branch is a cover term for a number of different dialect groups and individual dialects. The Northwestern subgroup, which centers on the German-speaking language area, consists of the Finnish Romani and Sinti varieties. Exhibiting a number of shared innovations and strong German influence, closely related Sinti-Manuš varieties are thought to have evolved in German-speaking areas, whence certain Sinto communities migrated further to the likes of the Netherlands, Austria, Hungary, Slovakia, Russia, and northern Italy. The Finnish Romani variety is closely related to Sinti, with a few thousand native speakers in Finland. Within the Northern Branch Matras proposes a Northeastern Polish-Baltic-Russian subgroup, with the North Russian or Xaladitka dialect, the closely related dialect of Polska Roma in Central Poland, and Latvian Romani spoken by a small population in Lithuania and Estonia.

A separate branch of the Northern Romani dialect group was British Romani, which was displaced by the majority language by the end of the 19th century and survives today only in the form of a special lexicon, as a Para-Romani variety (for the discussion of the term and concept see: Matras, 1998b). Also in this branch was Welsh Romani, which still had native speakers in the mid-20th century, but today survives only in documents such as Sampson's 1926 grammar. Based on its linguistic features, Iberian Romani may also have belonged to the Northern group, but these dialects are similarly extinct, with traces surviving in Spanish- and Basque-based Para-Romani varieties (Bakker, 1995) known as 'Calo'.

As dialect classification test cases the author mentions isolated and difficult-to-classify dialects, such as Abruzzian and Calabrian Romani, and Croatian/Slovenian Romani varieties documented in Slovenia and Italy.

Chapter 3 is devoted to the discussion of the historical and linguistic origin of the Romani-speaking populations. Through a study of the layers of the basic inherited lexicon and historical analysis of phonological and morphological changes, the chapter presents the processes leading to the formation of Proto-Romani and Early Romani. Proto-Romani, the pre-European "ancestor" of Romani, can be reconstructed by comparison of other Indo-Aryan proto-languages, the modern Indo-Aryan languages evolving from them, and present-day Romani dialects. Proto-Romani is thought to have separated from subcontinental Indo-Aryan during the transition period between Middle and New Indo-Aryan. Based on historical and linguistic data the predecessors of the Rom are believed to have inhabited Asia Minor by as early as the 11th century. Persian and Armenian elements in Romani lead historical linguists to believe the ancestors of the Rom lived in the Near East prior to the Byzantine period. The evolution of Early Romani is placed at the Byzantine period, an era marked by intense and long-lasting contact with Greek, the influence of which still appears in the structure of all present-day Romani dialects. Greek had a strong impact not only on the Romani lexicon (approx. 220–250 Greek-derived items), but also on the derivational and inflectional morphology as well as the morpho-syntax. Linguistic contact with Greek is attributed as the motivation of the emergence of a preposed definite article, the Verb-Object word order, and the split between factual and non-factual complementisers. The Greek component indicates that Early Romani was a member of the Balkan Sprachbund in the late Byzantine period. The end of the Early Romani period came in the 13th and 14th centuries with migration from the Balkans to inner Europe, the strengthening of divergent developmental tendencies and the split into dialect branches.

Chapter 4 (Descriptive Phonology) is devoted to the discussion of Romani phonological system. In the Romani consonant set the stops are stable elements: The three inherited Indo-Aryan stops, labial /p/, dental /t/ and velar /k/ are retained in all dialects. The author distinguishes three consonant groups based on palatalization as an articulation feature: genuine palatal stops, palatalized consonants and affricates. The genuine palatals are recent developments in the language. Both the genuine palatals and palatalized consonants may emerge internally, relying on internal palatalization of dentals in selected lexemes, or may be triggered through contact as well. As for affricates, both the aspirated and non-aspirated voiceless postalveolar affricates merge in a palatalized affricate /č'/.

Sonorants proved to be a less stable consonant group in Romani, because they are often subject to substitution or metathesis. The basic sonorants of Romani are a labial nasal /m/, a dental nasal /n/, a dental lateral /l/ and a dental trill /r/. The lateral /l/ is partly velarised in most dialects. Early Romani probably had a sonorant /ř/, which became /l/ in some Balkan Romani dialects. This historical /ř/ is continued as a uvular /R/ in Kelderaš and a number of Balkan dialects, or as a long or geminate trill /rr/ in various dialects. The substitution of the trill by a uvular /R/ in the Romani varieties spoken by the Sinti and Manuš is a result of German and French influence. Elsewhere it has merged with /r/.

In the Early Romani phase the fricatives in the language were probably the labials /f, v/, a velar /x/, a glottal /h/, dental /s, z/ and postalveolar sibilants /š, ž/. Velar /x/ was probably the result of a Proto-Romani innovation (/kh/ > /x/). Some dialects also have palatal sibilants /ś, ź/, resulting from reduction of the affricates /čh, dž/. Semi-vowels had a relatively marginal role; /j/ occurs primarily as a result of morpho-phonological jotation and in prothetic position. Early Romani affricates probably included postalveolar /č, dž/, an

aspirated postalveolar /čh/ and a dental /c/. As a result of innovations in certain dialects, both the number of affricates and their frequency of occurrence increased.

In Romani consonants are also characterized by voiced/voiceless opposition. Aspiration in Romani is a phonological feature that indicates its nature as a New Indo-Aryan language. There are aspirated counterparts for all three major stops /ph, th, kh/, as well as for the dental-postalveolar affricate /čh/. The latter is the weakest element of the Romani set of aspirates, often being subject to phonological changes. Consonant gemination may also be considered a new phenomenon, and is largely restricted to certain dialects.

The Early Romani vowel inventory probably consisted of the basic five-vowel system, /a, e, i, o, u/. The appearance of other vowel qualities (e.g., centralized /e/ and /i/) is a contact phenomenon. Vowel lengthening is an areal contact feature. Vowel length is independent of stress. The main stress pattern in Romani is the word-level grammatical stress. In the inherited pre-European (thematic) component of Romani the stress falls on the final position of lexical roots, in the absence of grammatical affixes, or on Layer I inflectional endings in nominal categories, on person inflection in finite verbs, and on the final component of indeclinables. Indeed, this pattern is also characteristic of athematic European components in cases where the borrowed elements are adapted to the Indo-Aryan inflectional morphology. Stress is documented to have moved forward in Romani varieties influenced by contact languages (e.g., Hungarian) where stress is word-initial, or others where it is penultimate.

The author analyzes phonological and morpho-phonological processes such as variations of the historical retroflex /nd/ in modern Romani dialects, prothesis and truncation in lexical roots, morphophonological jotation, *s/h* alternation in grammatical paradigms (Matras, 1999b), and the retention or loss of final *-s* in various dialects. Based on varying participation of certain dialects in these phonological changes, diagnostic isoglosses important in dialect classification can be drawn (see: Chapter 9).

Chapters 5 and 6 discuss questions of morphology. The morphological structure of Romani is equally characterized by inflective tendencies inherited from Older Indo-Aryan, agglutinative tendencies parallelling certain developments in Modern Indo-Aryan, and more recent analytical tendencies. In addition to these inherited morphological processes, the formation of the Romani morphological system is also influenced by grammatical borrowing from European contact languages.

Chapter 5 analyzes the morphology of nouns and nominal categories (cf. the studies in Elšik and Matras, 2000). The author treats gender, animacy and what is called ''thematic status'' as inherent properties of the Romani noun; that is, properties not expressed at the sentence level (cf. case) or discourse level (cf. definiteness). Romani distinguishes two genders, masculine and feminine. Gender is an important aspect in the noun inflection paradigm, and in the agreement of head noun and its modifiers. Gender agreement is also mandatory for pronouns and articles. The animate/inanimate distinction has a role mainly in marking direct objects. Inanimate nouns take the nominative, while animate nouns receive an oblique marker. Individual dialects may also show animacy split with other case markings and with pronominal reference. The most unique property of the Romani noun is its thematic status.

The terms of thematic versus athematic status and grammar are used primarily in English-language literature on Romani linguistics. The terms refer to a morphological split

which systematically differentiates European borrowings from inherited elements of the pre-European lexicon, treating them differently. The term 'thematic' refers to morphological patterns of predominantly Indo-Aryan origin which are used for the pre-European lexicon and a few early European (mainly Greek) loans. In contrast, Romani lexemes dating from the European period and borrowed from Greek and other, later contact languages receive different, ''athematic'' endings.

In Romani linguistics the thematic versus athematic opposition refers to the above mentioned morphological distinction between inherited and borrowed vocabulary which is present throughout the Romani morphological system. Historically this distinction evolved in the Early Romani period, tracing back to the adoption of a number of Greek morphological endings. For example, morphemes of Greek origin are used in the nominative of certain nouns, in loan verbs as adaptation affixes, and in adjectives as derivational and inflectional suffixes. Comparative linguistics indicates that the Greek morphemes borrowed by Romani for the purpose of loanword integration had a similar function in the Greek language itself, for which reason Romani athematic morphology is often referred to as the borrowing of a Greek borrowing pattern (Bakker, 1997). The thematicity split is characteristic of all Romani dialects, although there are dialectal variations with regard to the form of the athematic elements. Thus, variation is found in the distribution of the Greek-derived nominal morphology, while later contact languages also contributed to the formation of the athematic morphology.

This chapter surveys the main processes of noun and adjective formation, with a detailed analysis of nominal inflection and the case system. The Romani case system, like that of other New Indo-Aryan languages, distinguishes three different layers (cf. Matras, 1997 for the details). Layer I elements comprise nominative and oblique suffixes which are connected directly to the nominal base and are usually stressed. The distinctions of gender, number and thematicity are also on this layer. The masculine singular and plural oblique endings (-es and -en) function as general oblique markers and are situated between the nominal base and Layer II case endings. In Romani the oblique form also occurs independently, in a number of varying functions. The Layer II category consists a closed, stable set of agglutinative markers for various cases (dative, locative, ablative, instrumental and genitive). Layer III contains analytic adpositions, the set of which is more open and changing. The author also provides a detailed analysis of adjective inflection and the pronoun system, with particular attention to the deictic and demonstrative elements.

Romani verb morphology is characterized predominantly by suffixation, with the exception of aspect-marking prefixes borrowed or calqued from Slavic languages, the use of which is observed only in a few dialects. Chapter 6 analyzes the basic verb structure both synchronically and diachronically, the main features being as follows. The core of the verb is the lexical root. Borrowed verbal roots are followed by loan adaptation affixes. Derivational extensions are typically attached to the lexical root or to the adapted loan root. These derivational extensions primarily mark valency alteration, as well as aspectual issues such as intensity or iterativity. Perfective aspect is marked by an extension (-d-, -l-, -t-) to the verb stem. Romani verbs are inflected for person and number by two different sets of personal endings: one of the sets accompanies the present stems, the other one the perfective stems. A rare typological feature is that the expression of time is external to aspect and person inflection. The remote tense is expressed by an agglutinating

suffix *-as/-a/-e/-s/-ys/-ahi* marking the pluperfect and imperfect. For marking the future research documents consistent dialectal variation. Some dialects exhibit an analytical tendency: future constructions contain an auxiliary (e.g., *l-* lit. 'take') or other particles (e.g., *ka* from the verb *kam-* 'want'). The expression of conditional and quotative modality similarly uses various particles (*te*, *bi*, *li*).

Matras treats transitive versus intransitive derivations as one of the basic dichotomies of the Romani verb system. This dichotomy is interpreted as the result of morphological operations participating in the process of valency alteration. Older Romani grammars discussed this phenomenon under the heading of verb formation, or as one of the classification feature of verb conjugation groups. Valency alteration has proved to be an especially productive process in dialects which are or were in intensive contact with agglutinating languages. (The process has been documented, for example, for Sepeči, a member of the Balkan dialect group in contact with Turkish, and for the Vlax variety Agia Varvara in Athens, as well as Central varieties in contact with Hungarian, such as Romungro and East Slovak Romani.) A number of productive synthetic morphological devices are used in various dialects for these purposes. The semantic and syntactic functions of the derivations generated are quite varied. For example, intransitive inchoative verbs to express change of state may be formed from nouns, adjectives or adverbs; e.g., *bar-(j)o(v)-* 'to grow' from *baro* 'big.' Intransitive derivations from verbal roots can render either synthetic passives (e.g., *kin-d-(j)o(v)-* 'to be sold' from *kin-* 'to sell') or lexicalized intransitives. Reflexive and reciprocal forms may also be derived from intransitive verb roots. Transitive verbs deriving from nouns and adjectives are referred to as factitives: e.g., *bar-ar-* 'to raise, to grow,' *gilj-av-* 'to sing' (<*gili* 'song'). Another important group of transitive verbs derived from verb roots is causatives (cf. Hübschmannová and Bubenik, 1997). The chapter also gives a com-prehensive analysis of the historical evolution of valency alteration markers.

Also studied in detail are the morphological devices for loan-verb adaptation, stem formation, the characteristics of present and perfective verb stems and inflection classes, and person concord markers. Unlike earlier grammatical descriptions, the author analyzes tense-aspect-modality categories in Romani in a unified, reinterpreted framework. Three dimensions are distinguished in Romani TAM categories: an aspectual dimension (perfective:non-perfective), a temporal one to distinguish categories of remote:non-remote, and a modal dimension for specifying morphological features relating to "intentionality" (e.g., subjunctive). Matras stresses that past time is not an inherent quality of the perfective. Perfective aspect is described as viewing the action or event it marks as one that has been completed prior to or at the contextual point of reference. For this reason the extended verb stems containing the markers -d-, -l-, or -t- – which the majority of earlier descriptions referred to as Preterite or Aorist – are interpreted as expressions of aspect, not tense.

The only modal form in Romani which is inherently non-indicative is the subjunctive. A number of configurations of time and aspect categories can also be created with the non-factual complementiser *te* for expressing various non-indicative modal meanings. The subjunctive typically occurs in clauses of non-factual semantics and optative constructions.

Chapter 7 is devoted to the questions of syntactic typology. First noun phrase structure is analyzed. In Romani the first slot in the noun phrase is reserved for prepositions. The next

position is reserved for the class of determiners: demonstratives, interrogatives, possessive adjectives, and definite articles. These determiners assign definiteness and are incompatible with one another. The genitive adnominal, acting a determiner, often belongs to this prenominal slot. In the linear order of prenominal positions the next slot is for quantifiers, and after that for adjectives. Matras further distinguishes a postnominal pragmatic 'option' position, occupied by attributive elements.

With regard to the word order of the verb phrase Matras identifies a 'conservative' VO word order type. Object fronting is a pragmatic option, occurring mainly in contrastive structures. This word order also applies for pronominal objects. Pronominal direct objects precede pronominal indirect objects. Neither SV nor VS word order dominates in Romani; the position of the subject alternates between these two types. Essentially the same word order rules apply to main clauses and subordinate clauses. In Romani there is no special word order for questions.

Examining the rules for the SV versus VS alternation, the chapter analyzes the relationship between word order and discourse-pragmatic motivations, as well as word order rules relating to predicate and subject type and to formal constraints. The chapter shows that SV order is usually interpreted as focused or topicalized. It analyzes the function of the two alternatives in terms of continuity versus discontinuity. VS is interpreted as connective-integrative-narrative (see Matras, 1995b), and SV as perspective-establishing word order. VS is characteristic when the subject is non-animate, indefinite or a non-agent. VS is most common in presentative constructions; sentences where the predicate is a verb of being or motion. Conditional or non-factual complement clauses introduced by the conjunction *te* are predominantly VS. In these clauses the verb comes immediately after the conjunction *te*.

The chapter also analyzes genitive structure, the various types of complex sentences, and negation. Matras points out that Romani relative clauses are of the European type in that there is no specialized connecting element (relativiser or relative pronoun) exclusively for relative clauses; instead, interrogatives are used as relativisers. The most common relativiser in Romani is *kaj*, from the interrogative *kaj* 'where.' There are two relativisers which are inflected: *savo* 'which' and *kon* 'who.' Both inflect for case, the former also agrees with its head in number and gender. If the head noun is not the subject of the relative clause, Romani requires a resumptive pronoun. The use of the resumptive pronoun is conditioned by hierarchy of animacy and thematic role, and more generally by the predictability of the semantic case role of the particular head noun in the relative case. When discussing complementation, purpose clauses and adverbial subordination, Matras focuses on the factual versus non-factual dichotomy, a characteristic Balkanism (Friedman, 1985) in Romani. There is a factuality distinction in the use of subordinators. The complementiser *kaj* is used in structures where the events are presented as potentially independent, real or factual. Modal complements are introduced by *te*. As modal subordinator or non-factual complementiser, *te* serves the semantic/pragmatic function of making the truth value of the predicate relative. The subordinator *te* is used mainly in various conditional clauses.

Chapter 9 deals with dialectal diversity and proposes a new model of dialect classification. In his discussion of the theoretical and methodological problems, Matras emphasizes that the proposed dialect classification is not a genealogical categorization system. Thus, the individual dialect groups are not to be understood as separate branches of

the language family-tree model used in traditional historical linguistics, but as linguistic geographical categories describing language change in a geographic diffusion model[2]. He also stresses that the placement of certain dialects in this system is relative, a question not of "absolute group membership" but of dialects close or not so close to one another in a geographical dialect-continuum.

This model depicts the relationship of individual dialect groups of Romani through isoglosses drawn on the basis of distribution patterns of different variants of diagnostic linguistic variables. The system presented here classifies Romani dialects primarily according to their participation in changes affecting the inherited structure. Matras points out that from the inventories of shared features it appears that the isoglosses that are diagnostic of one branch may be irrelevant to the definition of another. Consequently, Matras recommends an examination of diversification from the perspective of historical emergence and geographical diffusion of isoglosses. Based on a systematic comparison of dialects, different processes of change may be assumed. An example of this is the diffusion of innovations. Some changes are simplifications, while others are cases of option selection among inherited variation. The geographical distribution of different forms may also provide information as to the relative chronology of some historical changes.

Taking into account the number of shared isoglosses and their structural importance, the dialect groups were identified along with the central area from which certain changes are likely to have diffused. With regard to internal innovations Matras distinguishes three centers of diffusion. The first center is located in Southeastern Europe, whence innovations diffused to the northwest. (Changes spreading from the dialects of this region include the loss of final consonant in the inherited abstract nominal marker *-ipen > -ibe*, or the assimilation of stems with historical perfective markers in -*tj*- to the inflection class in -*lj*- or -*j*-.)

The second center of diffusion is located in western-central Europe, with innovations typically spreading to the east or southeast. Some of the innovations spreading from here: the relative loss of the active participle in favor of the person-inflected 3SG form, initial jotation in third-person pronouns (*jov* SG.M, *joj* SG.F, *jon* PL., gender-neutral) and in the word for 'egg', and the truncation of *a*- in words such as *amal > mal* 'friend', *akana > kana* 'now'. On the basis of the relevant data it appears that the center of the *a*-truncation isogloss is the Sinti. Other examples of the diffusion of innovations from the northwest to the southeast: reduction of the cluster **nd > *ndř* to *r* (cf. **mandřo > maro* 'bread', **pindro > piro* 'foot', etc.), and the full affrication of the dental in the interrogative *keti > keci/kici* 'how much.' The extension of *s > h* selection to interrogatives has its center in the Sinti group of Central Europe, which is often referred to in the literature as *h*- dialects because of its high degree of participation in *s > h* change.

The third geographical diffusion center is the Vlax dialect group, with innovations spreading south, north and westwards. Innovations spreading from here are: affrication (through palatalization) of the initial dental in selected items (e.g., *tikno > cikno* 'small'), the prothesis of *a*- in *nav > anav* 'name', the specialization of long forms of the present

[2] For the different concepts of language change and variability in dialectology and sociolinguistics, methods and basic concepts in dialectology and related criticism in sociolinguistics literature (see e.g., Chambers and Trudgill, 1980; Wardhaugh, 1992).

conjugation in -*a* for future tense, and the reduplicated demonstrative stems k_d_, and k_k_.

The author also lists a number of regional innovations. During the presentation of other classificatory features the author discusses option selection (e.g., the alternation of *s* and *h* in grammatical paradigm, stem extension -*in*-, subjunctive -*av*-/-*ov*- in the copula, etc.), simplification (e.g., reduction of genitive marker, simplification of demonstratives and the loan-verb adaptation markers) and retention of archaisms (e.g., retention of definite articles in -*l*-). Paradigm simplification and restructuring proved a good indicator of internal coherence in dialect groups.

Based on the geographic diffusion model of language change there appears to be a northwest-southeast division within the European Romani dialects. The Vlax group seems especially innovative and dynamic. Highly coherent Sinti varieties in the northwest comprise a second influential center of language change. This group is the launch point of a number of changes, as well as functioning as a border inhibiting further spread of innovations from the southeast. Finnish Romani exhibits patterns similar to those of the Sinti group, on the basis of which literature speaks of a relatively unified Northwestern Finnish-Sinti group. The Polish-Balkan-Russian Romani dialects in the northeast display a pattern similar that of the Finnish-Sinti group. The author presents the evolution of Romani dialects as a dynamic process taking place between these two centers. From this perspective the Central dialects were wedged between these two centers, and appear to have been equally exposed to influences from the Vlax and Northern dialects. The Central dialects are divided by a number of isoglosses, yet in spite of their differences they show great consistency in morphological structure: future tense conjugation, case markers, person concord markers, demonstrative pronouns and loan-verb adaptation; thus the Central group may also be considered coherent. The Balkan group show a greater degree of internal diversity, but the differences are systematic with regard to a number of diagnostic isoglosses.

In addition to internal developments, dialect classification also takes into account contact phenomena resulting contact with various (former and present) surrounding languages. As shown in Chapter 8, contact-related dialect differences primarily characterize the lexicon and morphosyntax, and may generally be traced to differences between the source languages. At the same time, the rules for the morphological adaptation of borrowed elements are rather uniform (cf. athematic grammar), for which reason contact phenomena may be considered secondary in dialect classification.

The relevance of older, contact-focused dialect classifications is undermined not only by the continuous process of language change but also by past (and present) changes in social and linguistic contact conditions resulting from migration or geopolitical changes. As the author emphasizes in Chapter 8, in many communities a difference should be made between ''old L2,'' contact languages which once had a prolonged strong influence on the Romani variety in question but are no longer spoken at all in the community, ''recent L2,'' languages now known only by the older generation if at all, and ''current L2'' contact languages.

A more refined picture of internal diversity in a language can be gained by supplementing the external, analytical perspective with the perspective of the native speakers themselves, then reinterpreting the data in light of the latter context. This also

reveals the methodological limitations of dialectology survey in data collection and analysis. The different variants of the relevant variables are often interpreted by the members of the Romani speech communities as not primarily or not at all a product of geographic distribution, assigning them instead with social meaning and identity-marking function. Of course, it varies which variables are considered socially indexical in different Romani communities, and even within a particular community may depend on social and interactional context. The description of language diversity and change in a geographical dialectological framework is insufficient to capture these interrelationships — in large part precisely because of the characteristics of the data collection methods it employs.

The last two chapters of the book discuss issues of language use. Chapter 10 analyzes general aspects of the sociolinguistic situation of Romani-speaking communities. It stresses that Romani is a non-territorial language, characterized by discontinuous spread of its speakers, thus its use cannot be tied to a contiguous, geographically or politically delineated area. Romani-speaking communities are in a unique socio-linguistic situation: the ubiquitous situation of the ethnic minority which makes it indispensable for them to acquire and use the language(s) of the host society, thus becoming bi- or multilingual. Given the importance of Romani-paired bilingualism, surprisingly few studies to date have dealt with the various aspects of individual and societal bilingualism. Few studies have been written on code switching, the language choice in an interaction, and research on language shift and language maintenance is also badly needed.

From the works cited in this chapter it can be seen that with regard to the relationship between Romani and surrounding languages the literature thus far has largely made do with the generalized view which classifies Romani-paired bilingualism as an example of Fishman's model (1967) of ''bilingualism with diglossia[3].'' On this I would make a few observations, bearing in mind the tasks and perspectives of future research.

Bilingual competence and language use varies by community, and even within a single community the pattern is often quite varied; without a thorough examination of regularities of language choice and use in the various domains, and of language attitudes and ideologies of the members of the community, only a superficial picture can be seen. If Romani-paired bilingualism is viewed as generally diglossic, it is assumed to be characterized by function distribution between the two languages, where the majority language is the high variety, the code selected in formal public situations, while Romani functions as the low variety used only in informal private situations such as communication within the family.

For social and historical reasons deriving from the peculiar characteristics of always being a minority, it is indeed true that Romani speakers did not and for the most part still do not have the opportunity to use their native language in domains typically dominated by the majority language (e.g., institutionalized education, state administration, science, or the media) or on topics related thereto. As a result, Romani currently possesses limited resources for serving these functions. This situation is a linguistic consequence of social dominance and power asymmetry influencing the corpus, status, and

---

[3] For analysis and criticism of the Fishman's and other broad theories of diglossia (see e.g., Fasold, 1993; Hudson, 2003; Schiffman, 1996) and some of the studies of the thematic issue (no. 157) of International Journal of the Sociology of Language (Focus on Diglossia), published in 2002.

circumstances of acquisition of the language which is often typical of the majority-minority relationship.

The fact that there is no possibility of using Romani in interethnic communication in the aforementioned public domains, does not mean that there are absolutely no intra-ethnic domains, rules and means (e.g., styles and registers) for formal public use of Romani in its speech communities (cf. Stewart, 1989, 1997; Kovalcsik, 1993). Thus it cannot be asserted that Romani speakers always use their native language as the low variety, predominantly to serve basilectal functions, or that they consider it a lower-prestige code than the majority language (cf. Kovalcsik, 1999). Obviously, there are clearly a number of Romani communities where language shift is at an advanced stage, where the above arrangement would be true; however, to present this as the general situation for Romani-paired bilingualism regardless of social context would be an overgeneralization and simplification. Consequently, the statement that Romani is generally used in basilectal functions is in need of refinement and thorough (re)contextualization.

Given the above, descriptions of Romani-paired bilingualism and the distribution of function between languages need to employ a more dynamic approach, one more sensitive to social and linguistic contexts and variability therein. This would require a shift in analytical perspective and terminology where formal public use of Romani within a Romani-speaking community or between the members of different Romani communities, as well as the use of majority languages by Roma in formal situations of interethnic communication are conceptualized not as an inherently hierarchical relationship, but each in its own context, taking into account the attitudes and language ideologies of the participants themselves.

Like any living language, Romani has many different varieties. Thanks to recent research, there is now substantial information on regional and contact varieties. Considerably less attention has been paid to language use in its social context in various Romani-speaking communities, sociolinguistic research currently being an under-represented area in Romani linguistics. Information concerning language use is often published as the 'by-product' of anthropological or ethnographic research.

Little is known on variability of Romani language use, styles, registers, the characteristics of different genres, rules of interaction in various types of speech situations, discourse patterns, and the like. Presumably there are methodological reasons for this, related to data collection techniques. In order to analyze the connection of the various language forms with interactional context and social structure, it is essential to examine community interactions and native language ideologies related thereto. This information requires longitudinal research based on participant observation and employing ethnographic methods which draws language data from the study of the speaking as an action embedded in social context[4]. The use of anthropological methods and perspectives in Romani linguistic fieldwork and data analysis may also contribute to greater emphasis being placed on the participants' own voices and discourses. Thus would be particularly important with regard to linguistic minorities like Romani-speaking communities, where

---

[4] For an anthropological linguistic criticism of the ''field linguistic'' perspective and data collection techniques as well as the distinction of ''native speaker as informant, consultant vs. social actor'' (see Duranti, 1994: 14–27, 1997: 98).

everyday thinking and sometimes even scientific discourse have formulated countless misconceptions and deficit theories regarding the language and vernacular (i.e., Romani-medium) culture.

Chapter 11 analyzes questions of status and standardization. Romani today is still primarily an oral language. Romani literacy (Matras, 1999c) is a relatively new phenomenon; consequently one cannot speak of normative written traditions, or established regional or international standard varieties of Romani. This fact obviously relates to the characteristics of the minority situation of the speech communities as described above, including the virtually total absence of legal and institutional support for Romani language use in community-external domains (Szalai, 1999) which characterized the majority of European states as late as the 1990's. Matras presents the main tendencies in language planning with regard to Romani-speaking communities in various European states, and analyzes recent processes which may contribute to Romani language use becoming established in new settings. At present it appears that a number of different local or regional standard Romani varieties (cf. Puxon, 1979; Hübschmannová, 1995; Friedman, 1995, 1997, 1999), are taking shape, although there is also a demand for the creation of a uniform, internationally codified superposed standard variety (cf. Kenrick, 1996; Courthiade, 1989, 1992). New fields of international cooperation between Roma living in different states and speaking different Romani varieties and second languages (for example, various forms of written and oral communication in European-level politics, and the appearance of Romani mailing lists and web-sites on the Internet) have similarly strengthened the trend for pluralistic, decentral codification. Communication in Romani on international forums is characterized by dialectal flexibility, basic-level orthographic compatibility, and the avoidance of L2-specific loan elements, these being replaced by Romani-internal "soft neologisms," or by widely recognized international lexemes. For this reason Matras believes that new forms of communication in these domains free of normative constraints may effectively contribute to the formation of a modern "common written Romani."

5. Concluding remarks

As the author admits in the introduction (p. 4), "Linguistics cannot undo social injustice," but what it can do is replace stereotypes with information and facts, or at least contribute to the questioning of stereotypes. Yaron Matras' research in Romani linguistics, including *Romani: A Linguistic Introduction*, beyond its substantial scientific merit is also an important contribution to the scientific tradition challenging linguicism and misrepresentations relating to Romani.

References

Bakker, P., Cortiade, M. (Eds.), 1991. In the margin of Romani. Gypsy languages in contact. Institute for General Linguistics, Amsterdam.

Bakker, P., 1995. Notes on the genesis of Caló and other Iberian Para-Romani varieties. In: Matras, Y. (Ed.), Romani in Contact. The History, Structure and Sociology of a Language. John Benjamins, Amsterdam, pp. 125–150.

Bakker, P., 1997. Athematic morphology in Romani: the borrowing of a borrowing pattern. In: Matras, Y., Bakker, P., Kyuchukov, H. (Eds.), The Tipology and Dialectology of Romani. John Benjamins, Amsterdam, pp. 1–22.

Bakker, P., Matras, Y., 1997. Introduction. In: Matras, Y., Bakker, P., Kyuchukov, H. (Eds.), The Tipology and Dialectology of Romani. John Benjamins, Amsterdam, pp. vii–xxx.

Bakker, P., Matras, Y. (Eds.), 2004. Bibliography of Modern Romani Linguistics: Including a Guide to Romani Linguistics. John Benjamins, Amsterdam.

Boretzky, N., 2000. South Balkan II as a Romani dialect branch: Bugurdži, Drindari, and Kalajdži. Romani Studies, ser 5 (10), 105–183.

Chambers, J.K., Trudgill, P., 1980. Dialectology. Cambridge University Press, Cambridge.

Courthiade, M., 1989. La langue Romani (Tsigane): évolution, standardisation, unification, réforme. In: Fodor, I., Hagege, C. (Eds.), Language Reform. History and Future. Buske, Hamburg, pp. 79–109.

Courthiade, M., 1992. Research and action group on Romani linguistics. Interface 8, 4–11.

Duranti, A., 1994. From Grammar to Politics. University of California Press, Los Angeles.

Duranti, A., 1997. Linguistic Anthropology. Cambridge University Press, Cambridge.

Elšik, V., Hübschmannová, M., Šebková, H., 1999. The Southern Central (ahi-imperfect) Romani Dialects of Slovakia and Northern Hungary. In: Halwachs, D.W., Menz, F. (Eds.), Die Sprache der Roma: Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext. Drava, Klagenfurt, pp. 277–390.

Elšik, V., Matras, Y. (Eds.), 2000. Grammatical Relations in Romani: The Noun Phrase. John Benjamins, Amsterdam.

Fasold, R., 1993. Diglossia. In: The Sociolinguistics of Society, Blackwell, Oxford, pp. 34–60.

Fishman, J.A., 1967. Bilingualism with and without Diglossia, Diglossia with and without Bilingualism. The Journal of Social Issues 2, 29–38.

Friedman, V.A., 1985. Balkan Romani modality and other Balkan languages. Folia Slavica 7, 381–389.

Friedman, V.A., 1995. Romani standardization and status in the Republic of Macedonia. In: Matras, Y. (Ed.), Romani in Contact. The History, Structure and Sociology of a Language. John Benjamins, Amsterdam, pp. 177–189.

Friedman, V.A., 1997. Linguistic form and content in the Romani-language press of the Republic of Macedonia. In: Matras, Y., Bakker, P., Kyuchukov, H. (Eds.), pp. 183–198.

Friedman, V.A., 1999. The Romani language in the Republic of Macedonia: status, usage and sociolinguistic perspectives. Acta Linguistica Hungarica 46, 317–339.

Halwasch, D.W., 1998. Amaro vakeripe Roman hi – Unsere Sprache ist Roman: Texte, Glossar und Grammatik der burgenländischen Romani-Variante. Drava, Klagenfurt.

Hudson, A., 2003. Toward the systematic study of diglossia. In: Paulston, C.B., Tucker, G.R. (Eds.), Sociolinguisitcs. Blackwell, Oxford, pp. 367–377.

Hübschmannová, M., 1995. Trial and error in written Romani on the pages of Romani periodicals. In: Matras, Y. (Ed.), Romani in Contact. The History, Structure and Sociology of a Language. John Benjamins, Amsterdam, pp. 189–205.

Hübschmannová, M., Bubenik, V., 1997. Causatives in Slovak and Hungarian Romani. In: Matras, Y., Bakker, P., Kyuchukov, H. (Eds.), The Tipology and Dialectology of Romani. John Benjamins, Amsterdam, pp. 133–146.

Kenrick, D., 1996. Romani literacy at the crossroads. International Journal of the Sociology of Language 119, 109–123.

Kovalcsik, K., 1993. Men's and women's storytelling in a Hungarian Vlach Gypsy Community. Journal of the Gypsy Lore Society ser 5 (3), 1–20.

Kovalcsik, K., 1999. Aspects of language ideology in a Transylvanian Vlach Gypsy Community. Acta Linguistica Hungarica 46, 269–288.

Matras, Y., 1994a. Structural Balkanisms in Romani. In: Reiter, N., Hinrisch, Leeuwen-Turnocová, (Eds.), Sprachlicher Standard und Substandard in Südosteuropa und Osteuropa. Harrasowitz, Berlin/Wiesbaden, pp. 195–210.

Matras, Y., 1994b. Untersuchungen zu Grammatik und Diskurs des Romanes. In: Dialekt der Kelderaša/Lovara, Harrasowitz, Wiesbaden.

Matras, Y. (Ed.), 1995. Romani in Contact. The History, Structure and Sociology of a Language. John Benjamins, Amsterdam.

Matras, Y., 1995b. Connective (VS) word order in Romani. Sprachtipologie und Universalienforschung 49, 189–203.

Matras, Y., 1997. The typology of case relations and case layer distribution in Romani. In: Matras, Y., Bakker, P., Kyuchukov, H. (Eds.), The Tipology and Dialectology of Romani. John Benjamins, Amsterdam, pp. 61–93.

Matras, Y., Bakker, P., Kyuchukov, H. (Eds.), 1997. The Tipology and Dialectology of Romani. John Benjamins, Amsterdam.

Matras, Y. (Ed.), 1998a. The Romani Element in Non-Standard Speech. Harrasowitz, Wiesbaden.

Matras, Y., 1998b. Para-Romani revisited. In: Matras, Y. (Ed.), The Romani element in non-standard speech. Harrasowitz, Wiesbaden, pp. 1–27.

Matras, Y., 1999a. Johann Rüdiger and the study of Romani in eighteen-century Germany. Journal of the Gypsy Lore Society, ser 5 (9), 89–106.

Matras, Y., 1999b. S/h alternation in Romani: an historical and functional interpretation. Grazer Linguistische Studien 51, 99–129.

Matras, Y., 1999c. Writing Romani: the pragmatics of codification in a stateless language. Applied Linguistics 20, 481–502.

Puxon, G., 1979. Romanes and language policy in Jugoslavia. International Journal of the Sociology of Language 19, 83–91.

Schiffman, H.F., 1996. Diglossia as a sociolinguistic situation. In: Coulmas, F. (Ed.), The Handbook of Sociolinguistics. Blackwell, Oxford, pp. 204–215.

Stewart, M.S., 1989. True speech: song and the moral order of a Hungarian Vlach Gypsy community. MAN 24, 79–101.

Stewart, M.S., 1997. The Time of the Gypsies. Westview Press, Oxford.

Szalai, A., 1999. Linguistic human rights problems among Romani and Boyash Speakers in Hungary with special attention to education. In: Kontra, M., Phillipson, R., Skutnabb-Kangas, T., Várady, T. (Eds.), Language: A Right and a Resource. Approaching Linguistic Human Rights. CEU Press, Budapest, pp. 297–315.

Wardhaugh, R., 1992. An Introduction to Sociolinguistics. Blackwell, Oxford.

Andrea Szalai
Research Institute for Linguistics, Benczur u. 33,
1068 Budapest, Hungary
E-mail address: szalaia@nytud.hu.