TWO HUNDRED YEARS OF FINNISH ROMANI LINGUISTICS[1]

Kimmo Granqvist

ABSTRACT

In Finnish Romani linguistics, a historical standpoint was accentuated for a long time. The historical-comparative paradigm still had a strong impact on studies of Finnish Romani in the latter half of the 20th century. Early Finnish Romani linguistics is furthermore characterized by concretism, lack of theoretical ambitions, and emphasis on data. Word lists and dictionaries are still available in intervals of a few decades from the late 18th century up to the present day. At the Research Institute for the Languages of Finland, much emphasis was put on lexicography and data collection during the 1990s, according to the scientific ideals that then prevailed in Fennistic studies. At the beginning the 21st century, the central themes of research shifted to phonetics, phonology and morphosyntax in accordance with the interests and training of linguists. Still, very few papers have been published on language sociology and sociolinguistics. Active language planning of Finnish Romani and strong efforts to develop a written standard for the language were initiated at the turn of the 1970s. Nowadays the most central language-planning organ is the Romani language board.

1. INTRODUCTION

Since the latter half of the 20th century, at least three relatively homogenous periods can be distinguished in Finnish Romani linguistics. The shifts between these periods have occurred less because of scientific revolutions in the sense of Kuhn, but rather because old actors have been replaced by new ones. The resources of Romani Linguistics have always been extremely limited in Finland,

---

[1] The introduction of this paper is based on my earlier paper, *Advances in Finnish Romani Linguistics: Researchers and Paradigms* (Granqvist 2009). Much of sections 2–5 build upon the introduction of *Suomen romanin äänne- ja muotorakenne* (Granqvist 2007) and my earlier paper *(Un)wanted institutionalization: the case of Finnish Romani* (Granqvist 2006). The discussion on language planning is partly based on my paper *Romanikielen huollon pulmia ja peruskysymyksiä* (Granqvist 2008).

so much of the work has undertaken by researchers with divergent backgrounds. Until now, Romani has never been an academic subject, and its researchers have applied methods adopted either from international research traditions of Romani, or from other disciplines.

For a long time, a historical standpoint was accentuated in Romani linguistics, emphasizing the relationship of Romani with Old Indo-Aryan and Middle Indo-Aryan languages, and showing a lesser interest in the synchrony of the language. The historical-comparative paradigm that had been established in Europe during the 19th and early 20th century had a strong impact on studies of Finnish Romani still in the latter half of the 20th century. The most prominent name of Finnish Romani linguistics was Pertti Valtonen, who had studied Indo-European linguistics at the University of Helsinki. His main works, a licentiate thesis (1968) and a PhD thesis that unfortunately remained unfinished, were dominated by the same kind of methodology that Sampson (1926) used in his grammar and dictionary of the Romani of Wales.

In emulation of Sampson (1926), Valtonen's main accounts on Finnish Romani were syntheses that attempted to cover all levels of language from phonetics to lexicon, but emphasizing morphology. Valtonen could never reach the same level of detail as Sampson did. At the same time, the scope of Valtonen's studies was quite different from contemporary Fennistic works that tended to specialize in certain, often peripheral, problems such as individual features of dialects including certain groups of lexical items and certain morphosyntactic features or sounds.

Common with the Fennistic paradigm were what Karlsson (1975: 181) calls "concretism" and the lack of theoretical ambitions. Concretism refers to a perspective that emphasizes the primacy of genuine data as a research object. As a concept, it deviates radically from both the structuralist and generative frameworks that were prevalent in linguistics at that time. Valtonen's works were not greatly influenced by the trends of contemporary linguistics: rather the linguistic structures were presented using simple means similar to traditional grammar books, and the attempts to form theory were very limited.

The emphasis on data probably stems from the *Junggrammatiker*. In the Fennistic tradition, Karlsson (1975: 183) explains it as an effort at concretism and positivism. In Finnish Romani linguistics, another motivating factor could be the earlier wide interest of amateurs and professional linguistics in particular in the lexicon of Finnish Romani, which resulted in abundant data that were suitable especially for the needs of historical linguistics, but not so much for the study of synchrony. Word lists and dictionaries are still available at intervals of a few decades from the late 18th century up to the present day. All kinds of data were accepted as materials for research: old word lists, spiritual texts, and tape-recordings, of which the earliest ones originate from the 1960s. The openness of

the researchers to all kinds of data is of course understandable as gaining access to data is extremely laborious for extralinguistic reasons.

The research of Finnish Romani is most strongly institutionalized at the Research Institute for the Languages of Finland. The Institute is still the only institution in Finland with permanent resources allocated to the research and language planning of Romani. The work on Romani has a history of 28 years. On 9 October 1981, the Ministry of Education appointed a working group to debate the issues of the Romani language on the initiative of the Advisory Committee on Romany Affairs. In its memorandum of 5 March 1982, the working group suggested that the data and learning material project should be subordinate to the Ministry of Education, and the administration of the project should be linked to the Research Institute for the Languages of Finland. Officially, the research and language planning of Finnish Romani have belonged to the tasks of the Research Institute for the Languages of Finland since 1996 when legislative changes were made.

At the Research Institute for the Languages of Finland, much emphasis was put on lexicography and data collection during the 1990s. The areas of emphasis were first defined by people with a Fennistic background according to the scientific ideals that prevailed in Fennistic studies at the time so that they were also well suited to the highly material-oriented profile of the Institute. The data collection included obtaining and saving as computerized corpora a number of old written materials, but also collecting and transcribing a spoken language corpus. Numerous resources were allocated for several decades towards the compilation of two dictionaries (Koivisto 1994; 2001). Furthermore, a word list and a reverse lexicon (Granqvist 1997; 1998) were published as a part of the lexicographic work.

The connection of Finnish Romani linguistics with the Fennistic tradition was broken at the beginning the 21st century when new, linguistically trained actors took over responsibility for the work. One could claim the core linguistic study of Finnish Romani using modern methods was initiated at that time. Finnish Romani has been studied at least by Lars Borin (2000), Pia Brandt-Taskinen (2001), Henry Hedman (2003; 2004), Viljo Koivisto (2001; 2002; 2003), Anna Maria Viljanen (2002, together with Kimmo Granqvist), Anette Åkerlund (2004), and Tuula Åkerlund (2002), but the research has been in fact mostly undertaken by two actors: Helena Pirttisaari (Univ. of Helsinki) and Kimmo Granqvist.

The central themes of research shifted to phonetics, phonology and morphosyntax in accordance with the interests and training of linguists (following quite closely the paradigm adopted at the Dept. of General Linguistics). The study of syntax has recently been initiated at the Research Institute for the Languages of Finland in terms of a four-year project that aims at producing a thorough description of the syntactic structure of Finnish Romani within a theoretical

framework that combines generative and variationist traditions. Corpora still play a significant role in the current research work, but theory formation gained an important position as a part of its scientific goal. Results have been published as large monographs, providing exhaustive analyses of Finnish Romani and a number of papers dealing with more specialized issues. The connection with the old tradition of Finnish Romani linguistics was not completely lost, as both the historism and relationship to the data remain almost the same.

2. LEXICOGRAPHY AND ETYMOLOGY

The history of Finnish Romani linguistics dates back to the end of the 18th century. Kristfrid Ganander (1741–1790), who was a chaplain of Rantsila, compiled a dictionary draft for Finnish Romani, and a 46-page essay on the Finnish Roma and their language. The draft of the dictionary is lost, but the manuscript of Ganander's essay *Undersökning om De så kallade TATTERE eller Zigeuner, Cingari, Bohemiens, Deras Härkomst, Lefnadsätt, språk m.m. Samt om, när och hwarest några satt sig ner i Swerige?* is preserved in the archive of the Swedish Royal Academy of Letters, History and Antiquities. The Academy used to organize annually an essay competition on poetry, eloquency and history. In December 1779, the Roma, their origins and language were selected as the theme of the competition. The title page of Ganander's essay shows that his manuscript arrived at the Academy on 4 August 1780. The essay comprises 34 paragraphs, a majority of which discuss the history and ethnography of the Roma. Linguistically invaluable are paragraphs 22–32, which are the first notes on the language spoken by the Finnish Roma, no more than 200 years after the arrival of the first Roma to Finland was documented. In the essay was included a 150-word glossary and nearly 60 phrases. Ganander's language material is considered reliable by Joki (1956: 22) and Valtonen (1968: 19), even though his transcription is occasionally unsystematic. I do not enter here into details concerning the language material itself, since it has been extensively studied by Etzler (1944), Joki (1956) and most thoroughly by Valtonen (1968: 16–27).

Adolf Ivar Arwidsson (1791–1858) was a historian, university teacher, writer and journalist, who is better known for his striving to improve the position of the Finnish language at the University of Turku. He also took notes on the language of the Roma he had met in 1817 in Padasjoki, where he was born. Arwidsson's notes were published by Bugge (1858) as a part of his paper *Vermischtes aus der Sprache der Zigeuner*, which was included in the yearbook *Beiträge zur vergleichenden Sprachforschung*. Arwidsson's notes were more concise than those of Ganander but they comprised fairly complete declension and conjugation examples in addition to lexical material (for a more detailed account, see Valtonen 1968: 28–32). K.J. Kemell (1805–1832) compiled a glossary of Finnish Romani. After his death, the

glossary was burnt as an ungodly work, but Thesleff (see below) later created his own dictionary (1901) based on a draft of Kemell's glossary. According to Valtonen (1968: 33), this draft ended up with Professor Otto Donner via an auction, but it has since been lost. Jürgensen and Schmidt provided data on the lexicon of Finnish Romani to Miklosich who published them; these data were also included in Thesleff's (1901) dictionary.

Henrik August Reinholm (1819–1883) was actually an archaeologist and folklorist who made extensive notes on the Roma and their language when he was working as a prison preacher in Viapori and at a spinning house in Turku. Reinholm's notes are currently in two folders (number 87) titled *Finlands zigenare* at the Finnish National Museum. Folder 1 contains old publications, newspaper scraps and one sheet of paper written on both side. Folder 2 contains 892 hand-written pages (Sirkku Dölle, private communication, 19 January 2004). Reinholm's extensive but mixed data were compiled into a glossary containing grammatical notes on nominal and verbal inflections and approximately 2,000 words, which were also included in Thesleff's (1901) dictionary. A part of Reinholm's notes concerning the history of civilization were published by Aspelin in *Uusi Suometar* (1894–1895).

Arthur Thesleff (1871–1920) has been regarded as the most famous name within the study of the Finnish Roma (Valtonen 1968: 46). Because of his assets, Thesleff was elected president of the Gypsy Lore Society for three years in 1901. Most of his work was, however, ethnographical or sociological. His only accomplishment in Romani Linguistics remains his dictionary *Wörterbuch des Dialekts der finnländischen Zigeuner* (1901), which was published in the series *Acta Societatis Scientiarum Fennicae* and was the first printed dictionary of Finnish Romani. It concentrates mostly on the sub-dialect spoken in Western Finland and therefore contains numerous Germanic/Scandinavian loans but few borrowings from Finnish. It comprises 7,574 lemmas, but, according to Valtonen (1968: 48), the number of roots is about 2,100. The dictionary nevertheless covers almost the entire known lexicon of Finnish Romani, the size of which Thesleff (1901) himself estimated to be about 2,200 roots. Thesleff's estimation does not differ much from those published later (2,000–2,600 roots depending on the source).

It was said that Oskari Jalkio (Johansson) (1882–1952), who founded Suomen Mustalaislähetys (Romano Missio), was the only *gadžo* who mastered fluently the Finnish Romani language (Valtonen 1968: 51). Jalkio wrote articles and small books on Romani and the Finnish Roma, and translated songs and gospels into Romani. He also collected a glossary almost as extensive as Thesleff's dictionary. The manuscript of Jalkio's glossary is owned by Romano Missio. Based on Jalkio's data, Bourgeois (1911) wrote a small grammar of Finnish Romani called *L'esquisse d'une grammaire du romani finlandais*, intended for foreign students of

the language. Axel Kronqvist (1871–1956) compiled a dictionary of the western sub-dialect of Finnish Romani at the beginning of the 1950s. According to Valtonen (1968: 54) there existed two copies of the dictionary: one was owned by a private individual in Helsinki and the other by Romano Missio. A more recent glossary of Finnish Romani was collected by Yrjö Temo. His Finnish–Romani wordlist, containing 4,478 words, was donated to the Research Institute for the Languages of Finland in 1984.

The old glossaries of Finnish Romani, compiled by Ganander (1780), Arwidsson, Kemell, Reinholm, and others are invaluable for the study of the diachrony of Finnish Romani because they reveal many details of language change over intervals of just a few decades. The old glossaries are, on the other hand, incoherent, even within themselves, which results in uncertainty about what the researcher who took the notes meant. The orthographical practices vary considerably between the glossaries. The consonants [š, x, tš, dž] are denoted by the symbols *š*, *sh*, *sch*, *s*; *ȟ*, *ch*; *č*; *tš*, *tsch*, *tj*; and *dž*, *ds*, *dsch*, *ǧ*, *ž*, (*tš*, *tsh*), respectively. The graphemes *i* and *j* are interchanged, e.g. *mui* vs. *muj* 'mouth' (Granqvist 1997). Thesleff (1901) denotes long vowels using a macron, e.g. *bāro* 'big'. The second problem of the old glossaries is the large number of *hapax legomena* they contain (Granqvist 2001b). These *hapax legomena* are lexemes that may be neologisms invented by the compilers of the glossaries themselves or by the informants even purposefully to mislead the note-taking *gadžo*. Some of the *hapax legomena* may occur in the glossaries due to error. According to Grant (1995), the invented lexemes constitute a problem in many vocabularies of European Romani dialects.

Lexical items representing sub-dialects of Finnish Romani are included to a varying extent. While Thesleff's (1901) dictionary was mainly based on the Western sub-dialect, he included several variants from the Eastern [Karelian] sub-dialect, in which the sound change *š* >*x* has not taken place: *šēl* pro *xeel*, and certain lexemes show *ä* in front of *i* similar to some Russian Romani dialects, where *a* has been fronted into *e* due to a regressive assimilation triggered by the following high front vowel *i*, cf. Finnish Romani *tšai* ~ *tšäi* 'girl', Russ. Rom. *čhej* 'girl'. Even the lexical forms are different from glossary to glossary. Since Finnish Romani lacks a true infinitive, ACT.IND.PRES.1SG tends to be used as lexical entries for verbs, e.g. *rakkavaa* 'speak:PRES.1SG-REM', but Temo provides rather combinations of the complementizer *te* and a plural form of the subjunctive, e.g. *te rakkaven* 'COMPL speak:PRES.2/3PL'.

Pertti Valtonen published an etymological dictionary of Finnish Romani in 1972. Valtonen's etymological dictionary covers about 1,800 roots and provides numerous examples. In addition to being an etymological dictionary, it is actually one of the best general-use dictionaries of Finnish Romani. Pentti Aalto edited Sjögren's word list based on the language of the Roma of Ingria (Aalto 1982). Aalto

(1982: 31) regarded the Roma of Ingria as possible ancestors of the Lajenge Roma of Estonia, who were exterminated by the Nazi occupiers during World War II.

At the Research Institute for the languages of Finland, the main area of emphasis was lexicography until 2001. Most of the resources for Romani were allocated to assisting Viljo Koivisto in compiling the two dictionaries. Dictionary projects in general constitute an important part of the work conducted by the Research Institute for the Languages of Finland. Dictionaries are, furthermore, an early strategy in the standardization of languages that have lived mainly in oral form (Bernard 1996; Diarra 2003). For a long time there were no resources dedicated to Romani, but instead a number of researchers at the Institute, mainly Fennists, devoted some part of their time to Romani. Viljo Koivisto was assisted in the compilation of the Romani–Finnish–English dictionary by temporary personnel paid by the hour. Viljo Koivisto himself participated in the work through a monthly paid grant. Koivisto's (1994) Romani–Finnish–English dictionary comprises approximately 5,500 lexical entries, most of which are declined. Extremely valuable are its 5,800 examples that shed light on the use of the Romani lexical items. Viljo Koivisto's (2001) Finnish–Romani dictionary comprises about 23,000 lexical entries, a huge number of which are, however, more or less transparent collocations translated from Finnish compound words. The new dictionaries have provided a good basis for further research and material production. They do not suffer from the same drawbacks as the early word lists since internal coherence is better maintained and more criticism is applied in the selection of lexical items (cf. Granqvist 2001a). The Research Institute for the Languages of Finland has in addition published a few papers of lexicographical content, minor glossaries (Granqvist 1997; Jussila 1997), and a reverse lexicon of Finnish Romani (Granqvist 1999c).

3. DATA COLLECTION

The early interest in Finnish Romani manifested itself chiefly in collecting lexical items and providing notes on the grammar; the data were preserved in manuscript form. The academic interest in Romani increased in general in the 1960s and 1970s, at which point Pekka Sammallahti, Matti Leiwo and Pekka Jalkanen began to tape-record the language. Data collection was emphasized along with lexicographical work as a part of the activities of the Research Institute for the Languages of Finland during 1990s and the beginning of the 2000s. This work resulted in a number of corpora of both written and spoken Finnish Romani. The written language corpora include all the old glossaries, Thesleff's (1901) dictionary, gospel translations, Romani news presented by the national broadcasting company YLE; Viljo Koivisto's articles published in *Romano Boodos* were also obtained while producing the reader *Drabibosko liin* (Koivisto 2001).

The spoken Romani corpus has been transcribed from tape-recordings, partly done in 1995 by Mr. Juhani Pallonen (Research Institute for the Languages of Finland) during a Romani language seminar, and partly by Mrs. Hellevi Hedman-Valentin in 2000–2001. The transcription used is quite broad, as the aim is mainly to provide material for lexical and syntactic studies. The size of the resulting corpus is approximately 168,000 words. The number of informants is 89 (46 women and 43 men). The informants live in 32 municipalities in Finland, so that both sub-dialects of Finnish Romani are equally represented. The age of the informants varies between 16 and 87 years. The interviewers are teachers of Romani, with good skills in Romani. The interviewers used, however, a learned and quite formal code, which was practically monolingual (primarily Romani). While all interviewees have a relatively good proficiency in Finnish Romani, in the material there are several indications that Finnish is for them the better-known language.

Another corpus of Finnish Romani has been compiled by Lars Borin. Borin's corpus (approximately 110,000 words) consists mostly of written language: 170 articles published in *Romano Boodos,* Viljo Koivisto's (1982, 1987) text books, his translation of the gospel of John (1971) and the spiritual song book *Deulikaane tšambibi* (1970), the memorandum of the Orthography Committee *Mustalaiskielen normatiivi sanasto* (1971) and the transcriptions of Paul Ariste's (1940) interviews that he carried out in the 1920s. (Borin 2000)

4. STRUCTURE OF ROMANI AND LANGUAGE CONTACTS

While most of the early accounts on Finnish Romani aspired to describe the vocabulary, almost all of them also provided at least some information about the structure of Finnish Romani, such as declension patterns. The Estonian professor Paul Ariste conducted a few studies of Finnish Romani and showed that the extent of Finnish interference was more considerable than had been thought (e.g. 1940; 1966). The first academic theses on Finnish Romani were compiled in the 1960s by Pertti Valtonen at the Institute for Asian and African Studies at the University of Helsinki. His master's thesis discussed the Indo-Aryan words in Finnish Romani (Valtonen 1964) and his licentiate thesis the diachrony of Finnish Romani in the light of notes from different time periods and his own fieldwork (Valtonen 1968). Unfortunately he did not finish his PhD thesis *Indoarjalaisen perusaineksen säilyminen Suomen mustalaisten kielessä*: the manuscript is kept at the Institute for Asian and African Studies. Furthermore, he published in the *Journal of the Gypsy Lore Society* a collection of texts in Romani (Valtonen 1967) as well as an article describing the history of scholarly work (Valtonen 1969). His etymological dictionary (1972) is still an important source for Romani Linguistics. Matti Leiwo has also published a number of articles on

Romani phonology (1970) and language planning (1999). Pentti Aalto, who was a professor of Sanskrit and comparative Indo-European linguistics, studied the research history of Finnish Romani (Aalto 1977).

Except for Pertti Valtonen, none of the researchers really concentrated on Romani. One might think that Romani would have been institutionalized at the Institute for Asian and African Studies at the University of Helsinki as a part of its field of work, but in fact the interest in the language was personalized in Valtonen and slackened when he stepped away.

At the early 2000s, Romani gained a foothold at the Department of General Linguistics, which has long shown an interest in studying minority languages such as Sign language. In the early 2000s, two basic courses in Finnish Romani were arranged. By now the department has produced two master's theses in Romani. Pia Brandt-Taskinen (2001) discusses the verb complements of Finnish Romani and focuses on the question of whether Finnish Romani has an infinitive or a subjunctive. Helena Pirttisaari's (2002) master's thesis was a very comprehensive analysis of the participles in Finnish Romani, but it actually dealt with a much wider range of matters of verb morphology and lexicon. Her master's thesis is best characterized as truly revolutionary in recent Finnish Romani linguistics conceptually and methodologically and with regard to its extent. In her master's thesis, as well as in her later work (Pirttisaari 2002; 2003; 2004a; 2004b; 2004c; 2005) on Finnish Romani verbs, she adopted a functional-typological paradigm (Martinet 1962; Greenberg 1966; Anttila 1972; Coseriu 1974; Givón 1985a; 1985b; Haiman 1985). She was the first one to apply a structuralist model and Natural Morphology to Finnish Romani (Dressler 1977; 1985; Mayerthaler 1981; Wurzel 1984). Her master's thesis was the first account of Finnish Romani to discuss more thoroughly the partial loss of thematic/athematic dichotomy.

Pirttisaari has concentrated on understanding and explaining linguistic variation and structural changes in diachrony in accordance with her functional-typological starting point. Through her work, the study of the relationship between form and function (iconicity, isomorphy, economy, etc.) and complexity of language has become organic parts of the methodology of Finnish Romani linguistics. She has studied most exhaustively analogical changes in inflectional morphology as well as type and text frequencies as factors behind linguistic change. Some of her papers have followed methods known from Fennistic studies, such as Finnish field morphology and speaker profiles.

Granqvist's (1999a; 2001c; 2002b; 2004; 2005; 2007; Granqvist & Pirttisaari 2003; Granqvist & Hedman 2003) recent work has concentrated on describing Finnish Romani phonology and morphology. He has used the methods of experimental phonetics to shed light on the sound system and prosodic features. Granqvist has utilized models of nonlinear phonology such as autosegmental

phonology and feature geometry, and Optimality Theory. During 2001–2002 Kimmo Granqvist implemented Koskenniemi's (1983) Two Level Morphology for Finnish Romani using the PC-Kimmo environment. The version of PC-Kimmo Granqvist used contains a unification-based word grammar, due to which the parser is also suitable for syntactic analysis (Granqvist 2005). In Granqvist (2004; 2007), he applied Stump's (2001) and Spencer's (ms) Paradigm Function Morphology to Finnish Romani. PFM is a modern model of inferential-realizational morphology and represents Stem-and-Paradigm morphology. This work has resulted in a number of papers and a 739-page monograph that comprises a detailed account of Finnish Romani phonology and morphology as part of European Romani in its entirety. Granqvist's starting points include the emphasis of formalism and the strong position of UG in determining which grammatical categories are inflectional/derivational as well as maintaining a distinction between the grammar and the lexicon (Separation Hypothesis, Beard 1995). Descriptive economy is a central issue in Granqvist's work due to the many different rule types of PFM.

The monograph describing Finnish Romani phonology and morphology was published in 2007 (Granqvist 2007). The work leading to a syntax monograph described was initiated in 2007. The study of Romani syntax is conducted within the Minimalist Program (Chomsky 1993; 1995). Interesting theoretical issues are related to case stacking in secondary/oblique cases encoded by agglutinating suffixes of historically postpositional nature to oblique suffixes which derive from Old/Middle Indo-Aryan case markers, and to *suffixaufname* in the genitive.

Still much of our actual knowledge of the structure of Finnish Romani is based on teaching materials. The very first known attempts to provide learning materials in Romani took place as early as the late 19th century. The preface of A. Lindh's *Mustalaiskielinen aapinen* bears the date of 30 October 1893, but the book was never printed. The first printed ABC-book in Romani was Viljo Koivisto's (1982) *Drabibosko ta Rannibosko byrjiba*, which was followed five years later by a more demanding textbook called *Kaalengo tšimbako sikjibosko liin* (Koivisto 1987). In 1995 was published Miranda Vuolasranta's ABC-book *Romani tšimbako drom*, and in 1996, Henry Hedman's grammar guide *Sar me sikjavaa romanes*. *Drabibosko liin* (Koivisto 2002) is a reader in Romani that collates the author's earlier articles with a variety of themes. A majority of these had been previously published as articles in the newspaper *Romano Boodos*, which is issued by the child welfare organization Romano Missio. For elementary instruction in the Romani language, a workbook together with a teacher's guide was recently published (Vuolasranta et al. 2003).

5. SOCIOLINGUISTICS AND LANGUAGE SOCIOLOGY

Very few studies have been conducted outside the core linguistics that one might see as "normal science". Work outside normal science tends to result from occasional co-operation, efforts to participate in on-going international discussions, or themes of congresses and requests from publications. A few papers have been written or published on the language sociology of Finnish Roma (an unpublished work by Granqvist, ms.; Granqvist & Viljanen 2003), and on the institutionalization of Finnish Romani (Granqvist 2006), but these studies only complement the normal science. Linguistic taboos and politeness strategies constitute perhaps the most extensively studied area within language sociology. Currently, a master's thesis is being prepared at the University of Helsinki by Petra Kovanen on Romani-Finnish code-switching.

Raino Vehmas's PhD thesis *Suomen romaaniväestön ryhmäluonne ja akkulturoituminen* (1961) actually belonged to the field of sociology, but it contained some invaluable observations regarding the usage of the Romani language (Vehmas 1961: 91–99, 188–189). The data that Vehmas published were based on an earlier survey that was carried out by the Social Bureau of Research in 1954. According to statistics published by Vehmas (1961: 91–99, 188–189), 60% of adult Roma had a complete or good command of Romani, and as many as 89% considered themselves able to get along in Romani. On the other hand, Finnish was the main language of discourse for a vast majority of the informants. Already at that time, the proficiency of young Roma in Romani was weak, which Vehmas (1961: 188) interpreted as attributable either to a slow rate of language acquisition, or to an accelerated pace of acculturation. Another survey was made by the Social Welfare Office of Helsinki in 1979; in this survey, 185 Roma families were interviewed. This survey indicated that no more than 37% of the families that participated in the study mastered the language well, and only 21% could get along in every-day situations (Mustalaisasiain Neuvottelukunta 1981: 57–58). Kopsa-Schön (1996) discusses, *inter alia*, the use of Romani based on interviews of 127 Roma. According to her (1996: 44), only the elderly Roma were able to communicate fluently in Romani in the 1990s. The young Roma she interviewed did not speak Romani. The most thorough study of the position and use of Finnish Romani is being prepared by Henry Hedman. The study is based on interviews of 306 Finnish Roma in Finland (262 informants) and Sweden (44 informants). The informants were chosen to represent equally both sexes and all ages. Geographically, the interviews cover most areas where the Finnish Roma live.

6. LANGUAGE PLANNING

Active language planning of Finnish Romani and strong efforts to develop a written standard for the language were initiated at the turn of the 1970s when many activists began to realize the degree of attrition and loss of domains. In 1970, the Ministry of Education appointed an orthography committee whose task was to develop an easy-to-read orthography for Romani and to compile a normative lexicon (MNS, 1971). The chairman of the committee was Pentti Aalto, and the members were Viljo Koivisto, Pekka Sammallahti, and Mirjam Karimus. The orthography that was suggested by the committee is by now fairly well established in dictionaries, learning materials and some of the journal and newspapers articles, but certain characters cause technical problems.

Nowadays the most central language-planning organ is the Romani language board, which has acted as an expert body in connection with the Research Institute for the Languages of Finland since 1 June 1997. The language board has seven members, who are elected for a period of three years. Five of the members are representatives of the Romani population. The task of the language board is to make decisions-in-principle and recommendations regarding the use of Romani. The Romani language board primarily airs issues of written texts and public usage. The language board does not deal with spoken vernacular language, but it attempts to follow closely the practices of the spoken language in the codification, so that the standard would not become too difficult and distant for its users. Far-reaching historical reconstructions as well as adaptations to other Romani dialects are avoided so that the results would not be unfamiliar for today's speakers. Rather the language planning aims at clarity and understandability. (Leiwo 1999)

During its eleven years of activity the Romani language board has taken up mostly lexical and terminological issues and to some extent the structure of the language. One of the most critical matters remains creating terminology for the needs of the modern society. Attempts have been made to codify the numerals and expressions of time. On grammatical issues perhaps the hottest discussion was devoted to grammatical gender, which has been mostly lost in Finnish Romani, possibly due to the contact interference of Finnish. The language board decided to follow the practice of the spoken vernacular and recommended maintaining the gender opposition in such cases as are motivated by natural gender.

The multilingualism of the Roma and the dominance of Finnish have reflections in all language planning. Clarity and understandability of new expressions manifest themselves as an effort to retain syntactic-semantic transparency with the corresponding Finnish lexical items. Most neologisms are loan translations from Finnish, e.g. *folkengo naaluno* 'deputy', *bringosko ~ bukkosko goono* 'pleura' (Viljo Koivisto), *horttas ranniba* 'orthography' (Viljo Koivisto). The transparency helps the Roma understand texts in Romani, but is also connected with the

common means of the Roma to produce Romani by translating from Finnish. The effort to gain transparency may also cause structural changes in Romani. Dialects of European Romani do not favour compounding, but Roma have begun to treat collocations consisting of an adjective attribute or genitive adnominal and a noun orthographically as compounds (e.g. *aprunodikkiboskiero* 'supervisor', *arteskomuseos* 'arts museum') in order to facilitate the perception of the notions expressed by the collocations as a whole. Furthermore, the Finnish Roma show a tendency to form Finnish-like compounds consisting of two nouns in the nominative (*verboskaľit* 'verb tree') despite the fact that this type of word formation is generally unknown in Romani dialects.

Following the spoken vernacular is connected with retaining the clarity and understandability of the Romani language. One obvious disadvantage of following the spoken vernacular has been that considerable variation is permitted in the language. Partially, forms representing different sub-dialect are permitted: Finnish Romani is divided into a Eastern and Western sub-dialect approximately in the same way as Finnish (Granqvist 2002b; 2007). The written standard of Finnish Romani [to the extent the dialect has been codified until now; the codification is on-going] has received more influence from the Western sub-dialect than the Eastern one due to the seminal standardization and lexicography work done by Viljo Koivisto since the 1970s, but a majority of the members of the Romani language board is made up of speakers of the Eastern sub-dialect. Therefore the language has a tendency to accept parallel forms based on the Eastern sub-dialect. In practice, all flowers blossom. Producing "correct" Finnish Romani has been made easy for those who already master the language, because they do not need to memorize many new rules. On the other hand, teaching materials need to account for a huge amount of variation, which not only makes compiling them a laborious task, but also makes teaching or learning the normative language very difficult.

Linguistic purism is the opposite pole of following the spoken vernacular both in codification and revitalization of Romani. Extremely puristic attitudes have manifested themselves in particular in literary production and creating neologisms. In both cases, the most transparent loan words from the closest contact languages tend to be avoided (in particular those with a phonological form that bears a strong resemblance to Scandinavian or Finnish lexemes); neologisms are most readily based on Old Indo-Aryan and Middle Indo-Aryan linguistic material or material considered as such. But the thematic component, which has been reduced to 1,000 roots even in Europe, does not constitute a sufficient lexical resource. Already now extensive polysemy and tendencies to changes due to variation manifest themselves. Furthermore, the resulting collocations, usually consisting of recursive genitive adnominals, are structurally highly complex, cf. *barokomujengosikjiboskosentrum* 'big:mask + people:GEN.PL + teaching:GEN.SG + centre = centre for adult education'.

Characteristic of the language planning have been rather conservative or protective attitudes towards the development of Finnish Romani. There is little or no inclination at all to borrow lexical items or grammatical features from other Romani dialects. Lexemes are borrowed from English, Swedish or other contact languages rather than from other Romani dialects. In contrast, the Finnish Roma who live in Sweden and have tighter contacts with Vlax-speaking immigrant communities show opposite tendencies and seemingly attempt to adapt their Romani usage to gain mutual comprehensiveness. Finnish Romani is both peripheral and isolated among the Northern dialects and European Romani dialects in general, but the isolation probably has been intentional. Except for Sweden, the Finnish Roma have had little contact with other Romani communities in Europe. They have also shown little interest in learning other Romani dialects to be able to extend their possibilities of communicating with other Roma.

7. RESEARCH AND ROMANI AS A SECRET LANGUAGE

Finnish Romani lived for a long time solely as an oral language of the Roma community, used within the family and as a secret language (Valtonen 1968: 241). The language constituted a symbol of cultural identity and formed a boundary that separated the Roma from *gadže* (Åkerlund 2002: 126; Hedman 2004: 42–43), but it has also provided protection as well as an opportunity to discuss the family's internal matters in alien places (Åkerlund 2002: 127).

As the political climate of the latter half of the 20th century permitted a number of legislative measures that greatly improved the position of the Roma and other minorities, the institutional rights of Finnish Romani were extended and new linguistic domains emerged: what had previously been private and solely restricted to the Roma community now became public. The public use of Romani was seen to contradict the traditional point of view, according to which the language, which was solely the property of the Roma, should be carefully safeguarded from outsiders (Åkerlund 2002: 126; cf. Ganander 1780; Valtonen 1968: 241–245; Grönfors et al. 1997: 175; Pirttisaari 2002: 17–18).

This nature of the Romani language has had serious consequences for language teaching, production of language materials and research. The Roma have accepted the production of materials on the condition that the authors are Roma and the books are only sold to Roma. Occasionally even half-Roma have not qualified as authors. All kinds of materials still encounter resistance from some traditional-thinking Roma. The participation of *gadže* in Romani issues is a matter of controversy among the Roma. Because of this, most non-Roma researchers have met, at least to some degree, resistance, in particular on the part of Roma activists. During the first years of its activity, the Romani language board acquired a strong

role as an overseer of all the research (intended to be) conducted, even though this role was not actually included in its tasks: the language board required all research plans to be submitted to it and was able to obstruct the study of Finnish Romani grammar for several years. The language board still decides about researchers' permissions to use the language materials at the Research Institute for the Languages of Finland.

The attitudes towards *gadže* having an opportunity to learn Romani also hamper education. However, for teacher training, production of new, adequate teaching materials and the continuity of the research it would be extremely important to widen the academic interest for Romani.

8. DISCUSSION

During the last fifty years of Romani Linguistics in Finland, three chief paradigms, or ways of doing normal science can be distinguished: the historical-comparative work conducted by Indo-Europeanists, the material-oriented and lexicographic work preferred by researchers having a Fennistic background, and finally, the "linguistization" of the work as result of the participation of actors with a background in General Linguistics.

Along with what I wish to call "linguistization", the knowledge of Finnish Romani and number of publications and presented papers have increased dramatically. Since 2000, two master's theses have been prepared on Finnish Romani by linguistically trained scholars, one large monograph has been published, more than twenty scientific articles have been published and still others are forthcoming, and some fourty papers have been presented in national or international scientific forums.

The "linguistization" has skewed the division of resources at the expense of language planning and revitalization. Currently the permanent resources of Romani linguistics are located at the Research Institute for the Languages of Finland. Open discussion regarding the strategies and goals, more resources and new, competent actors will be required to be able to produce linguistically significant insights on the language and to apply the results in a manner that fully benefits the primary client of the work done on Finnish Romani, the Roma community. At the same time, the goals of the language planning need to be revised and the question of education to be solved.

ABBREVIATIONS

1	1st person
2	2nd person
3	3rd person
ACT	active
COMPL	complementizer
GEN	genitive
IND	indicative
PFM	Paradigm Function Morphology
PL	plural
PRES	present tense
REM	remoteness marker
SG	singular
UG	Universal Grammar

REFERENCES

Aalto, Pentti 1977. Zur Geschichte der Erforschung des finnländischen Zigeunerdialekts. In: Johann Knobloch & Inge Sudbrack (eds), *Zigeunerkundliche Forschungen*, I. (Innsbrucker Beiträge zur Kulturwissenschaft, Sonderheft, 42): 61–72. Innsbruck: AMŒ.

Aalto, Pentti 1982. *Le vocabulaire du tzigane d'Ingrie composé par Sjögren*. (Études Finno-Ougriennes, XV) Akadémiai Kiadó, Budapest. Separatum.

Åkerlund, Anette 2004. Romaninuorten kieli Suomessa. In: Marja Nenonen (ed.): 237–239.

Åkerlund, Tuula 2002. Oman kielen merkitys romanien kulttuuri-identiteetille. In: Sirkka Laihiala-Kankainen, Sari Pietikäinen & Hannele Dufva (eds): 126–129. Anttila, Raimo 1972. *Historical and Comparative Linguistics*. (Current Issues in Linguistic Theory, 6) Amsterdam: Benjamins.

Ariste, Paul 1940. Über die Sprache der finnischen Zigeuner. *Õpetatud Eesti Seltsi Aastaraamat, Annales Litterarum Societatis Esthonicae* 1938/2: 206–221.

Ariste, Paul 1966. Mida on teada mustlasest ja nende keelest? *Küsimused ja vastused* n:o 6.

Beard, Robert 1995 *Lexeme Morpheme Base Morphology*. Stony Brook, NY: SUNY Press.

Bernard, H.R. 1996. Language preservation and publishing. In N. Hornberger (ed.), *Indigenous literacies in the Americas*. Mouton de Gruyter.

Borin, Lars 2000. A corpus of written Finnish Romani texts. In: Donncha O Croinin (ed.), *LREC 2000. Workshop proceedings. Developing language resources for minority languages: reusability and strategic priorities*: 75–82. Athens: ELRA.

Brandt-Taskinen, Pia 2001. Suomen romanikielen verbikomplementit. Master's thesis. University of Helsinki.

BUGGE, Sophus 1858. Vermischtes aus der Sprache der Zigeuner. In: A. KUHN & A. SCHLEICHER (eds), *Beiträge zur vergleichenden Sprachforschung auf dem Gebiete der arischen, celtischen und slawischen Sprachen*, 1: 143–155. Berlin: Dümmlers.

CHOMSKY, Noam 1993. *Language and Thought*. (Anshen Transdisciplinary Lectureships in Art, Science and the Philosophy of Culture) London: Moyer Bell.

CHOMSKY, Noam 1995. *The Minimalist Program*. Cambridge, MA: The MIT Press.

COSERIU, Eugenio 1974. *Synchronie, Diachronie und Geschichte*. (Internationale Bibliothek für allgemeine Linguistic, 3) München: Wilhelm Fink.

DIARRA, E. 2003. Choice and description of national languages with regard to their utility in literacy and education in Angola. In: A. OUANE (ed.), *Toward a multilingual culture of education*: 333–348. Hamburg: UIE.

DRESSLER, W.U. 1977. *Grundfragen der Morphonologie*. Wien: Österreichische Akademie der Wissenschaften.

DRESSLER, W.U. 1985. *Morphonology: The Dynamics of Derivation*. Ann Arbor: Karoma Publishers.

ETZLER, Allan 1944. *Zigenarna och deras avkomlingar i Sverige. Historia och språk*. (Stockholm studies in Scandinavian Philology, 4) Uppsala: Almqvist & Wiksells Boktryckeri.

GANANDER, Kristfrid 1780. Undersökning om De så kallade TATTERE eller Zigeuner, Cingari, Bohemiens, Deras härkomst, Lefnadssätt, språk m.m. Samt om, när och hwarest några satt sig ner i Swerige? Käsikirjoite. Tukholma: Vitterhetsakademien.

GIVÓN, Talmy 1985a. Language, Function and Typology. *Journal of Literary Semantics* XIV/2: 83–97.

GIVÓN, Talmy 1985b. Iconicity, isomorphism and non-arbitrary coding in syntax. In: John HAIMAN (ed.), *Iconicity in Syntax*. Amsterdam: Benjamins.

GRANQVIST, Kimmo 1997. *Suomi–romani-sanaluettelo*. Helsinki. Moniste.

GRANQVIST, Kimmo 1998. S*uomen romanikielen käänteissanasto. Reverse Lexicon of Finnish Romani.* (Kotimaisten kielten tutkimuskeskuksen julkaisuja, 111) Helsinki.

GRANQVIST, Kimmo 1999a. Vowel Harmony in Finnish and Finnish Romani. *SKY Journal of Linguistics* 12: 27–44.

GRANQVIST, Kimmo 1999b. Notes on Finnish Romani Phonology. *Grazer Linguistische Studien* 51: 47–63.

GRANQVIST, Kimmo 2001a. Suomen romanikieli. *Kielikello* 1: 4–8.

GRANQVIST, Kimmo 2001b. Suomen romanikielen leksikografia ja sen erityisongelmat. *Leksikograafiaseminar Sõna tänapäeva maailmas. Leksikografinen seminaari Sanat nykymaailmassa*. (Eesti keele instituudi toimetised, 9) Tallinn: Eesti Keele Sihtasutus.

GRANQVIST, Kimmo 2001c. Suomen romanikielen puolivokaali /j/ ja puolivokaalin muodostukseen liittyvät morfofonologiset prosessit. *XXVIII Kielitieteen päivät Jyväskylässä 1.7–19.5.2001. Esitelmien tiivistelmät.*

GRANQVIST, Kimmo 2002b. Finnish Romani Phonology and Dialectology. *SKY Journal of Linguistics* 15: 61–83.

GRANQVIST, Kimmo 2004. Suomen romanin kestoilmiöiden fonologiaa. In: Marja NENONEN (ed.): 27–33.

Granqvist, Kimmo 2005. ROMTWOL. An implementation of a two-level morphological processor for Finnish Romani. *Proceedings of the 6*th *International Conference on Romani Linguistics*. München: Lincom Europa.

Granqvist, Kimmo 2006 (Un)wanted institutionalization: the case of Finnish Romani. *Romani Studies* 16: 43–62.

Granqvist, Kimmo 2007. *Suomen romanin äänne- ja muotorakenne*. (Suomen Itämaisen Seuran suomenkielisiä julkaisuja, 36; Kotimaisten kielten tutkimuskeskuksen julkaisuja, 145) Helsinki.

Granqvist, Kimmo 2008. Romanikielen huollon pulmia ja peruskysymyksiä. *Kielikello* 1.

Granqvist, Kimmo 2009. Advances in Finnish Romani Linguistics: Researchers and Paradigms. In: Jarmo Lainio, Annaliina Gynne & Raija Kangassalo (eds), *FinnForum VIII. An interdisciplinary Conference in Finnish, Finnish-North American and Sweden Finnish Studies.* (Reports from the Centre for Finnish Studies, 9): 131–138. Eskilstuna: Mälardalen University, Sweden.

Granqvist, Kimmo & Henry Hedman 2003. Suomen romanin nominimorfologiaa. *Virittäjä* 4: 482–507.

Granqvist, Kimmo & Helena Pirttisaari 2003. Suomen romanin leksikaalisista kerrostumista ja niiden fonologisista ja morfologisista prosesseista. *Puhe ja kieli* 23:1: 25–41.

Granqvist, Kimmo & Anna Maria Viljanen 2002. Kielelliset tabut romani-identiteetin kuvaajana. In: Sirkka Laihiala-Kankainen, Sari Pietikäinen & Hannele Dufva (eds): 109–125.

Grant, Anthony 1995. Plagiarism and lexical orphans in the European Romani lexicon. In: Yaron Matras (ed.), *Romani in contact: the history and sociology of a language*: 53–68. Amsterdam: Benjamins.

Greenberg, Joseph H. 1966. *Language universals, with special reference to feature hierarchies*. (Janua Linguarum, Series Minor, 59) The Hague: Mouton.

Grönfors, Martti, Kari Virolainen, Tuula Åkerlund & Paavo Lounela 1997. Suomen romaniväestö: unohdettu kulttuuriryhmä. In: Juha Pentikäinen & Marja Hiltunen (eds), *Suomen kulttuurivähemmistöt.* (Suomen Unesco-toimikunnan julkaisuja, 72): 149–184. Helsinki: Yliopistopaino.

Haiman, John 1985. *Natural Syntax*. Cambridge: CUP.

Hedman, Henry 1996. *Sar me sikjavaa romanes. Romanikielen kielioppiopas*. Jyväskylä: Opetushallitus.

Hedman, Henry 2001. *Luukasko Evankeliumos*. Jyväskylä: Suomen Pipliaseura.

Hedman, Henry 2004. Suomen romanikieli salakielenä. In: Marja Nenonen (ed.): 42–48.

Joki, Aulis 1956. Kristfrid Gananderin mustalaistutkielma. *Virittäjä* 1/1956: 1–27.

Jussila, Raimo 1997. Den finländska romanin och dess ordböcker. *LexicoNordica* 4:75–81.

Karlsson, Fred 1975. Fennistiikan tieteenparadigmasta ja sen ohjausvaikutuksesta. *Virittäjä* 79(2): 179–193.

Koivisto, Viljo 1970. *Deulikaane tšambibi. Hengellisiä lauluja*. Forssa: Mustalaislähetys.

Koivisto, Viljo 1971. *Johannesesko Evankeliumos*. Suomen Pipliaseura.

Koivisto, Viljo 1982. *Drabibosko ta rannibosko byrjiba*. Ammattikasvatushallitus − Kouluhallitus: Helsinki.

Koivisto, Viljo 1987. *Rakkavaha romanes*. *Kaalengo tšimbako sikjibosko liin*. Ammattikasvatushallitus − Valtion painatuskeskus: Helsinki.

Koivisto, Viljo 1992. Tutkimus Suomen romanikansan ammateista ja niissä tapahtuneista muutoksista. Master's thesis. Helsingin yliopisto.

Koivisto, Viljo 1994. *Romano-finitiko-angliko laavesko liin. Romani-suomi-englanti sanakirja. Romany-Finnish-English Dictionary*. (Kotimaisten kielten tutkimuskeskuksen julkaisuja, 74) Helsinki: Painatuskeskus.

Koivisto, Viljo 2001. *Suomi-romanisanakirja. Finitiko-romano laavesko liin*. (Suomalaisen Kirjallisuuden Seuran toimituksia 811, Kotimaisten kielten tutkimuskeskuksen julkaisuja, 117) Helsinki.

Koivisto, Viljo 2002. *Drabibosko liin*. (Kotimaisten kielten tutkimuskeskuksen julkaisuja, 120) Helsinki.

Koivisto, Viljo 2003. Katsaus romanikielen opetukseen 1970-luvulla Suomessa. *Latšo diives* 1: 18−20.

Kopsa-schön, Tuula 1996. *Kulttuuri-identiteetin jäljillä. Suomen romanien kulttuuri-identiteetistä 1980-luvun alussa*. Helsinki: SKS.

Koskenniemi, Kimmo 1983. *Two-level morphology: A general computational model for word-form recognition and production*. (Publication, 11) University of Helsinki, Department of General Linguistics, Helsinki.

Laihiala-Kankainen, Sirkka, Sari Pietikäinen & Hannele Dufva (eds) 2002. *Moniääninen Suomi*. Jyväskylä: Soveltavan kielentutkimuksen laitos.

Leiwo, Matti 1970. Äännesubstituutioista: erityisesti mustalaiskielen klusiilien mukautumisesta suomen fonologiseen järjestelmään. (Turun yliopiston fonetiikan laitoksen julkaisuja, 7): 41−50. Turku.

Leiwo, Matti 1999. Suomen romanikielen asemasta ja huollosta. In: Seppo Pekkola (ed.), *Sadanmiehet. Aarni Penttilän ja Ahti Rytkösen juhlakirja*. (Suomen kielen laitoksen julkaisuja, 41): 127−139. Jyväskylä: Jyväskylän yliopisto.

Martinet, A. 1962. *A functional view of language*. Oxford: OUP.

Mayethaler, Willi 1981. *Morphologische Natürlichkeit.* (Linguistische Forschungen, 23) Wiesbaden: Aula-Verlag.

Mustalaiskielen normatiivi sanasto. Eripainos Mustalaiskielen ortografiakomitean mietinnöstä 1971: A 27. Moniste. Helsinki.

Nenonen Marja (ed.) 2004. *Papers from the 30th Finnish Conference of Linguistics, Joensuu, May 15−16, 2003*: 27−33. Joensuu: University of Joensuu.

Pirttisaari, Helena 2002. Suomen romanin partisiippien morfologiaa. Master's Thesis. University of Helsinki.

Pirttisaari, Helena 2003. Muutos ja variaatio Suomen romanin verbien taivutustyypeissä. *Virittäjä* 4: 508−528.

Pirttisaari, Helena 2004a. Contact-induced changes in Finnish Romani. At: www.ling. helsinki.fi/kielitiede/20scl/abstracts.shtml\#Pirttisaari (accessed 9 Aug. 2004).

Pirttisaari, Helena 2004b. Variation and change in the verbal morphology of Finnish Romani. In: Marja Nenonen (ed.): 178–184.

Pirttisaari, Helena 2004c. Contact-induced changes in Finnish Romani: Passive Voice. At: www.ling.helsinki.fi/uhlcs/LENCA/LENCA-2/information/datei/29-pirttisaari. pdf (accessed 1 Jan. 2006).

Pirttisaari, Helena 2005. A functional approach to the distribution of participle suffixes in Finnish Romani. *Proceedings of the 6*th *International Conference on Romani Linguistics*. München: Lincom Europa.

Sampson, John 1926. *The Dialect of the Gypsies of Wales being the Older Form of British Romani Preserved in the Speech of the Clan of Abram Wood*. Oxford: Clarendon.

Spencer, Andrew ms. Generalized Paradigm Function Morphology.

Stump, Gregory 2001. *Paradigm Function Morphology. A Theory of Inflectional Form*. Cambridge: CUP.

Thesleff, Arthur 1901. *Wörterbuch des Dialekts der finnländischen Zigeuner*. (Acta Societatis Scientiarum Fennicae, XXIX:6) Helsinki: Finnische Litteratur-Gesellschaft.

Valtonen, Pertti 1964. Indoarjalaiset sanat Suomen mustalaisten kielessä. Master's thesis. Helsingin yliopisto.

Valtonen, Pertti 1966. Sanskriittiako, vaiko vain sekakieli? *Kotitiellä* 2: 44–47.

Valtonen, Pertti 1967. Finnish Gypsy Texts. Edited, translated and annotated, with the help of M. Valtonen, by B. Gilliat-Smith. *JGLS* XLVI/3–4:82–97.

Valtonen, Pertti 1968. Suomen mustalaiskielen kehitys eri aikoina tehtyjen muistiinpanojen valossa. Licentiate thesis. University of Helsinki.

Valtonen, Pertti 1969. Finnish Romani Research. *JGLS* XLVIII/3–4:124–127.

Valtonen, Pertti 1970. *Markusesko evankeliumos*. Helsinki: Kristillisen kirjallisuuden seura.

Valtonen, Pertti 1972. *Suomen mustalaiskielen etymologinen sanakirja*. (Tietolipas, 69) Helsinki: SKS.

Valtonen, Pertti 1979. Trends in Finnish Romani. In: Ian Hancock (ed.), *Romani sociolinguistics.* (International Journal of the Sociology of Language Series, 19): 121–125. The Hague: Mouton.

Valtonen, Pertti ms. Indoarjalaisen perusaineksen säilyminen Suomen mustalaisten kielessä.

Vehmas, Raino 1961. *Suomen romaniväestön ryhmäluonne ja akkulturoituminen*. (Turun yliopiston julkaisuja, sarja B, osa 81) Turku.

Vuolasranta, Miranda, Armas Hagert, Päivi Majaniemi & Henna Huttu 2003. *Romani tšimbako buttiako liin* 1. Helsinki: Opetushallitus.

Wurzel, Wolfgang 1984. *Flexionsmorphologie und Natürlichkeit. Ein Beitrag zur morphologischen Theoriebildung.* (Studia Grammatica, 21) Berlin: Akademie-Verlag.