The Slovak-and-Czech dialect of Romani
and its standardization

MILENA HÜBSCHMANNOVÁ
and JIŘÍ V. NEUSTUPNÝ

Abstract

Standardization of Romani is an inevitable process that is likely to be
initiated independently in various societies where the language is spoken.
This paper examines the situation in the case of former Czechoslovakia.
On the basis of the analysis of written texts it confirms that variation exists
and is evaluated by Romani speakers. Elaboration of the language is taking
place. The paper argues that standardization differs in early modern,
modern, and postmodern societies and that the old models of standardization
should not be used for Romani. In the postmodern context the unification
component of standardization recedes into the background. The standard
should be polycentric and competence in semicommunication should be
developed. There is an urgent need for further elaboration of the norm,
which may involve both purism and mixing. The symbolic function of the
standard should not affect its use in interaction, and the standard should
be developed to become the property of all rather than only of the Romani
middle class.

The Issue

Romani is the last European language largely unaffected by standardiza-
tion. The linguistic commission, created by the International Romani
Union at the Fourth World Romani Congress held in Warsaw in 1990,
commenced work toward the establishment of a "constructed" standard,
bridging virtually all Romani dialects (Hancock 1993a). Hancock has
suggested that such a standard be based on the Kalderash variety of the
Vlach dialect (Hancock 1993b).

While there may be a need for a constructed standard of this kind in
the case of international meetings of the Roma elites and as a symbol of
Romani identity, the issue we propose to raise in this paper is both less
impressive and more urgent. In some European societies Roma children arrive at school with no knowledge or only a limited knowledge of the language of instruction. In some countries, such as Slovakia, the lack of codification of Romani has been quoted as an argument against the introduction of Romani even as an auxiliary tool of instruction. At the same time, in most European settings where Romani is still alive, its very existence is threatened. It is imperative for its survival to become an instrument that can be widely used by the masses in a large number of situations in which it fails at present.

These problems, if not others, cannot wait until a “constructed” standard is created. Even should it be created and formally accepted soon, it may further complicate, rather than alleviate, the use of Romani in schools as a bridging language and the maintenance of the language in Roma communities.

Under these circumstances, the issue is how to handle standardization in a way that is more appropriate to the reality of the millions of the Roma. The issue may be different in different societies that have a Romani community. In this paper we shall deal particularly with the Slovak-and-Czech dialect of Romani.¹ The problem will be discussed from the point of view of a theory of language planning (Neustupný 1983; Jernudd and Neustupný 1987, 1991; Jernudd 1990; Neustupný 1994) that leaves the door open to full acknowledgement of the need for standardization, but not of the traditional kind.

Romani today

A few words about the contemporary situation of the Romani language may be appropriate. Romani is not an insignificant language. The overall number of the Roma has been estimated by some as high as seven or perhaps ten million people. Statistics cannot be relied on because in censuses large numbers of the Roma have traditionally appeared under other ethnic groups, both in view of a low degree of ethnic self-consciousness and because of discrimination against them. In 1985 the maximum number of the Roma in some European countries was estimated in the statistics of the EEC as follows:

<table>
<thead>
<tr>
<th>Country</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yugoslavia</td>
<td>900,000</td>
</tr>
<tr>
<td>Romania</td>
<td>800,000</td>
</tr>
<tr>
<td>Hungary</td>
<td>600,000</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>500,000</td>
</tr>
<tr>
<td>Spain</td>
<td>450,000</td>
</tr>
</tbody>
</table>
Czechoslovakia 400,000
France 300,000
Greece 120,000
Britain 110,000
Italy 90,000
West Germany 65,000
Holland, Portugal, Ireland, Switzerland, Sweden, Austria, 13,500
Finland

According to these statistics, the percentage of the Roma population was highest in Hungary (4.6%), Bulgaria (4.7%), Yugoslavia (3.5%), Romania (2.8%), and Czechoslovakia (2.6%). However, within Czechoslovakia the Roma were concentrated in its eastern part: in the Slovak Republic 4.8 percent of the population was of Roma origin, while in the Czech Republic the figure was only 1.4 percent. Czechoslovak Roma sources estimated that the total number of the Roma population was double the EEC figure.

The presence of the Roma in the USA has been documented by Hancock (1975, 1990), and the 1991 census demonstrated their undisputable presence in Australia. According to M.G. Clyne (personal communication) the census numbers should be considered highly underestimated, an interpretation supported by the fact that a lively radio program broadcast in Romani exists in Perth.

In Eastern and Central Europe, if not elsewhere, the Roma represent an explosive social issue, generated by the assimilationist policies of the former Soviet-block governments. These policies failed both in their attempt to erase the ethnic identity of the Roma and in raising their standard of living to that of the mainstream population. Apart from small groups of the Roma, their economic status remains very low, with most of them being unqualified laborers and, as such, out of work when the overall economic situation deteriorates.

One of the stereotypes about the Roma is that they are an itinerant population. While this may be true about some West European countries, and some individual Roma groups in Eastern Europe, most of the East European Roma have been sedentary for centuries, normally living at the outskirts of settlements of other ethnic groups or, more recently, in town and city areas abandoned by the mainstream population. In former Czechoslovakia, from which the data of this study derive, the nonurban settings are typical for Slovakia, while urbanized Roma settlements exist both in Slovakia and in the Czech Republic. It would be unrealistic not to accept that the incidence of criminality among the Roma is high. However, this phenomenon cannot simply be relegated to the traditional
lack of respect for ownership: the problems grow on features of social structure generated by the unsuccessful attempt at assimilation.

While in the traditional Roma settings the original language and culture are relatively well preserved, in urban locations a process of "deculturation" (Neustupný 1993) takes place. Among other cultural features, Romani is often lost without being replaced by the full version of the mainstream language. Of course, there is a considerable number of those who have assimilated and of Roma intellectuals, many of them with a positive attitude to the maintenance of the Romani culture and language. The Roma active in the International Romani Union are an example of those who can play an important positive role in improving the situation.

**Standardization: what it involves**

Within sociolinguistics, the issue of standardization has basically been discussed from two different points of view.

The first approach to standardization accepts the term "standard" in its everyday usage. Language standards are compared with standards in the area of measures such as foot, inch, or cubit (cf. Cooper 1989). Although Cooper does say that the analogy should not be stretched too far, attention is still centered on the "single, widely accepted norm which is felt appropriate with only minor modifications or variations for all purposes for which language is used" (Ferguson 1962). The emphasis is on uniformity of the norm, rather than its elaboration.

The second view of standardization stems from a different experience. In languages in which a unified norm already exists, uniformization is not the major issue. In English the term "standard" has been used to refer to such norms (cf. Joseph 1987), but in other European languages they have variously been referred to as the "literary language," "written language," "Gemeinsprache," or similar (Daneš 1987). Under these circumstances, emphasis moves to the elaboration of the norm. For example, the Prague School theory of the "standard language" (spisovný jazyk in Czech) emphasized features required of the unified norm: flexible stability and functional differentiation, achieved by intellectualization and automatization (Havránek 1963 [1932]; Garvin 1964).

Codification, which is the fixing of the standard in grammars, dictionaries, or literary and other canons, normally accompanies both uniformization and elaboration.

When the two different approaches to language standardization and codification are combined, we obtain a picture of standard languages as (a) possessing unified norms, which are (b) elaborated to serve all needs
of a modern society, and which are also (c) codified. When we say "possessing" we mean not only that the norm is available but also that it has been accepted and is being used. However, modern standards are also characterized by a number of other features. They also (d) are closely based on the language of at least a section of the society in question (even though they are not identified with this section only), (e) serve as symbols of modern nations, and (f) are the property of the intellectual class (or, in many Third World societies, actively support elites).

It is important to realize that standard languages in this sense are a product of the early modern and modern historical period, a fact on which most theoreticians of language standardization agree (cf. Bartsch 1987; Joseph 1987; Coulmas 1994). These standards do not necessarily coincide with similar phenomena within traditional societies, Western or non-Western. Latin in medieval Europe or Classical Chinese in the Far East presented fairly unified and highly elaborated and codified norms, ready to serve most needs of the respective societies. In this sense they can be described as standards. However, they had no base in the language of the communities concerned and did not function as symbols of modern socioeconomic integration, and their distribution within the societies they served was greatly limited. Milroy and Milroy (1985) have correctly pointed out that "it is best to speak of standardization as a historical process which — to a greater or lesser degree — is always in progress" (1985: 22) and Daneš (1987) has noted that the standard is a fuzzy concept with a firm centre but a large periphery. The premodern and modern standards are members of the same family but they warrant a separate analysis.

We believe that the concept of the standard is changing at present with the transition of societies from the modern to the postmodern stage. The current historical period places emphasis on accepting variation, which becomes the key slogan in all areas of culture. Uniformity of norms, or any uniformity at all, loses its former appeal. As the bumper sticker has it, "different is beautiful." This fully applies to ethnic diversity, which is not seen as a weapon of socioeconomic importance any more, and to regional differences. Under postmodern conditions language variation, regional or social, is seen as something that should be preserved rather than eradicated. Recent theoretical studies of language standardization such as Milroy and Milroy (1985), Crowley (1989), or Tollefson (1991) all emphasize this point. This postmodern attitude to variation is also reflected in Stary's insistence that the "interventionist" attitude of linguistics to language variation should be abandoned (Stary 1990).
When considering the issue of the standardization of Romani we must take into consideration this new situation. We speak of the end of this century and the beginning of the twenty-first. Although the wave of postmodernization has not fully affected all East European nations, it will arrive within a decade or two. Attitudes to the standard are already changing in the case of Czech. For example, nonstandard elements are tolerated in broadcasting, a large number of alternatives have been accepted in the official spelling, and there is a renewed interest in Czech diglossia (distinction between the standard and the common language). A broadening of this perspective cannot leave the issue of Romani unaffected.

Variation in Romani spoken in the Czech Republic

1. *Dialectal groups*

In the course of World War II two different groups of the Roma lived on the territory of Bohemia and Moravia (today the Czech Republic), which was in practice directly administered by Hitler’s Germany: Czech Roma and the Sinti (German Roma). Their total number can be estimated at 8000 people. These Roma were the subject of genocide and only approximately 600 people lived to see the end of the war. In Slovakia, which throughout the war was an independent state and ally of Germany, tough anti-Roma laws were adopted, but in practice the Roma escaped extermination. After the reconstitution of the Czechoslovak Republic in 1945, massive migration of the Roma from rural Slovakia to the industrial areas of Bohemia and Moravia took place. While in Slovakia (according to the 1970 census) approximately three-fourths of the Roma lived in “Gypsy settlements” in the country; in Czech territory the same percentage settled down in urban environments, dispersed among the matrix population. The total number of the Roma in the present day Czech Republic is estimated to be close to 300,000 people, while in the Slovak Republic the population is probably in the vicinity of 0.5 million.

The linguistic classification of the Roma dialects has been presented many times. Hancock (1975, 1988), referring to the work of Terrence Kaufman, distinguishes between the following groups:

1. Southern (Greek and Zargari);
2. Balkan (Romania, Bulgaria, Yugoslavia, Hungary, and Italy);
3. Northern (including the Central group, Sinti, Baltic, Scandinavian, and British groups);
4. Iberian; and
5. Vlach (originally itinerant in the Balkans and in Eastern and Central Europe).

The only scholar who has presented a linguistically sound picture of variation in the Romani language spoken on the territory of former Czechoslovakia is Jiří Lipa (1965). He distinguished four different groups: Czechoslovak (corresponding to the Central group under [3] above), Hungarian, German (Sinti), and Rumanian (Vlach, i.e. Lovari and Copper Smith or Kalderash Romani). Needless to say, the terminology used in reference to language variation by the Roma themselves differs from the descriptors used in this paper.

1. **Czechoslovak group.** This group includes Czech (i.e. Bohemian and Moravian) Romani, after the holocaust virtually extinct, and a variety that originally could have been called Slovak Romani but should more appropriately be referred to as Slovak-and-Czech Romani today. The reason is that this variety spread with the postwar migration of the Roma to the present Czech Republic and is the language spoken by the majority of the Roma both in contemporary Slovakia and in the Czech Republic. Although there are no exact statistics, it can be estimated that not less than three-quarters of the total number of the Roma speak this dialect. This population overflows to southern Poland and Western Ukraine. After World War II, Slovak-and-Czech Romani was the first language and the language of daily communication among the Czechoslovak Roma in all generations. Today, it has decayed considerably, in particular in urban settings, where it is being replaced by a pidginized form, strongly affected by Slovak or Czech. At the same time a trend to relearn the language, when necessary, and to develop it has appeared among the educated Roma (Hübschmannová 1979).

The Slovak-and-Czech variety of Romani has been used in 15 Romani publications (mostly bilingual, only five of them entirely in Romani) out of the total of 16 published on the territory of former Czechoslovakia (Hübschmannová 1991, i.p.).

2. **Hungarian group.** As far as numbers of speakers are concerned, this is the second largest group of the Roma; it is estimated that it comprises approximately 10 percent of the total Roma population on former Czechoslovak territory. The Hungarian dialects are distributed mainly in Hungary, Austria, and Slovenia, although it should be noted that other varieties of Romani are also used in these countries. Traditionally the Hungarian Roma have been more strongly assimilated than the Czechoslovak group. The Hungarian variety of Romani retains some
conservative features such as the intransitive–transitive–causative triplets for verbs.

3. German group (Sinti). On the territory of the Czech Republic this is a small group among those who survived World War II. Unlike the case of speakers of the Slovak-and-Czech dialect, speakers of Sinti keep their culture and language secret and do not welcome researchers. The language survives in older generations, whereas younger members of this group possess only a passive knowledge.

4. Vlach group. The Vlach group, comparable in former Czechoslovakia in its size to the Hungarian Roma, was itinerant until 1959 when it was forced by the government to settle down. The Czech Vlachi belong to the Lovari group, which is represented almost all over the world, but particularly in Hungary, Romania, and Poland.

The language of this group is very much alive, being the first language of children and subsequently used widely in daily life. In Hungary most Roma intellectuals are members of this group, and a number of literary works have been published in the dialect. In Slovakia the poet Jozsef Ravasz has printed his poetry in the journal Roma, published in Bratislava. Although described elsewhere (e.g. Pobožniak 1964; Choli Daroczi and Feyer 1988), in the Czech and Slovak Republic the variety has received little analytical attention.

2. Internal variation within Slovak-and-Czech Romani

Social variation. To non-Romani observers the Roma population often appears to be internally undifferentiated. This is not the case. The fact that soon after the “velvet revolution” of 1989, 44 different organizations, political, cultural, or other, were registered shows how diverse the Roma population is. It is surprising that within a relatively large population of 600,000–800,000 people information about different fajta (‘clan’) and the character of their relations is well preserved. The folk terminology uses the terms “friendly” and “unfriendly.” In 1993 the Roma living in the Prague suburb of Smichov refused to read the monthly Romano gendalos because it was published by Margita le Rapačiskeri čhaj (‘Margita, the daughter of Rapač’), who was a member of an “unfriendly” fajta.3

A social distinction of considerable importance is that between žuže Roma (‘clean Roma’) and the degeša (‘unclean’), the latter being eaters of dog and horse flesh. Until 1986 in Krompachy (Eastern Slovakia) there were two quite separate settlements: žuže Roma and the degeša.
However, normally the two groups live in adjoining localities and, stereotypes apart, do not differ linguistically.

**Linguistic variation.** The major dialectal division within the Slovak- and Czech group has traditionally been referred to as a division between the *Vichodňarsko duma* 'Eastern variety' and *Zapadňarsko duma* 'Western variety'.

1. The *Vichodňarsko* variety. The 1970 census showed that close to 50 percent of the entire Roma population in Czechoslovakia still lived in the relatively small area of Eastern Slovakia. They spoke the *Vichodňarsko* variety.

Within this variety we can distinguish several regional subvarieties, among which the most important ones are the Humenné-Michalovce ('Southern') subvariety and the Poprad-Prešov ('Northern') subvariety. There are other subvarieties with more limited distribution, such as those of Svidník or Betliah.

The Humenné-Michalovce subvariety is characterized by the following features:
- In accented syllables e and o are pronounced as ej and ou respectively.
- This feature is also present in the Lovari variety.
- Čh in some words is replaced by t'h (mat'hi for mačhi 'fly', mort'hi for morčhi 'skin').
- The verbal infinitive equals the 3rd person plural, rather than 3rd person singular (te keren for te kerel 'to do').
- On the lexical level this subvariety borrows mainly from Hungarian and Ukrainian.

The Poprad-Prešov subvariety can be identified on the basis of features such as the following:
- Palatalized d is pronounced as dž, rather than d' of the other varieties; t is palatalized as č, rather than t' (džives for d'ives 'day', bući for but'i 'work').
- The infinitive is the same as in Western varieties (i.e. equals the 3rd person singular).
- The lexicon includes loanwords from Polish and German, borrowed through Slovak varieties (*biglajzis* 'an iron', *blavajzis* 'pencil').

2. The *Zapadňarsko* variety. A number of features characterizes this variety. Some of them are the following:
- This variety lacks the palatalized l.
- H appears at the beginning of some words instead of Eastern s (*har* for *sar* 'how') and t instead of c (*tikno* for *cikno* 'small').
A number of grammatical and lexical features distinguish this variety from the Eastern varieties. These features will be further discussed in the following section.

In rural Slovakia the regional distribution of the two main varieties corresponds to the division between Eastern and Western Slovakia. In the Czech Republic, both the Zapadňarsko ('Western') and the Vichodňarsko ('Eastern') variety (together with the Hungarian and Vlach dialects) are represented in most Romani speech communities.

**Variation in two Romani texts.** Two Romani texts have been analyzed to establish what linguistic forms distinguish, in written Romani, the two principal varieties. Both texts were written by Bartoloměj Daniel in the Zapadňarsko variety. The Vichodňarsko version of the texts was established with the help of native speakers of this variety.

Bartoloměj Daniel was born in Šaštin in Western Slovakia and after completing his military service in 1948 studied to qualify for entrance to the university. He completed a degree in history and subsequently worked in archives, as a Roma activist, and, in the period 1973–1989, as a manual worker. Currently he is curator in the Museum of Roma Culture in Brno. The texts were published in the first issue of the journal Romano džaniben ('Romani studies'), which appeared in 1994.

The first text is entitled Miro dživipen ('My Life') and narrates briefly the life history of the author. The second text Kije historia Romengri ('On Romani history') is linguistically of a more demanding nature because it involves a number of special terms that were not yet available in Romani.

In the following sections only words that have different alternatives in the Eastern varieties will be given. Abbreviations: E = Eastern variety (Vichodňarsko duma); EH = Eastern variety, Humenné-Michalovce subvariety; EP = Eastern variety, Poprad-Prešov subvariety; no abbreviation = Western variety. The forms will be given in the order in which they appear in the text for the first time, and the figure in parentheses indicates how many times the word appears.

1. First text: My Life

   a. Words with different alternatives in E: angoder (3), E. anglunes 'before'; rod'ina, E. famel'ija 'family'; amoñis, E. kovinca 'anvil'; keribnaskero (4), E. but'akero 'hardworking, laborer'; zatopinel, E. thovel jag 'to heat'; odija (2), EH. od'a EP. odoj 'there'; sasikano, E. slugad'iko 'military'; ph(a)ravel, EH. phuterel, EP. phundravel 'to open'; cheroj, E. čang 'leg' (in EP cheroj means 'thigh'); adeso (2)/asso, E. ajso 'such'; džid'ovel, E. dživel 'to live'; chalovel, E. achal'ol 'to understand'; romane
The Slovak-and-Czech dialect of Romani and its standardization


b. Neologisms (the abbreviation BD indicates that the word was probably coined by the author; other words are still perceived as neologisms but are used by others as well):

BD. phuvkeriben ‘agriculture’; BD. phuvkeribnaskero ‘farmer’; BD. uče prengre ‘high boots’ (cf. E. ciracha); šelberša ‘century’; BD. elementarno sikhlaribni ‘elementary school’; BD. rajkani sikhlaribiči ‘middle school’; BD. mangipen ‘wish’ (in the press the same word is often used with the meaning ‘requirement’); sikhlardengero ‘teacher’s’; gend’i (3)/gindo ‘book’; BD. uči sikhlaribiči ‘university’; džanibnaskero (2) ‘scientific, academic’; BD. chalaro (3) ‘specialist’; jekhetanib (3) ‘association’; jekhetano ‘common, shared’; angalalipen ‘virtue, strong point’; avrithod’i ‘exhibition’.

c. Grammatical variation

- The demonstrative ada ‘that’ has only one form ada for both the masculine and the feminine (e.g. ada buti, E. adi buti ‘that work’).
- The case that means designation (‘going to…’) has the ending -u (E. -ate), as in Jablonicu, E. Jablonicate (a toponym).
- A different construction is used in the case of ‘you must go’: moste phires, E. mulšines te phirel.
- In the sentence jekhe čhoneske chudiñas e but’i pre obca ‘in a month he got work with the local council’ two grammatical variants occur: čhoneske ‘month, dative’, E. čoneha ‘month, instrumental’ and chudiñas, E. chudnās/chudl’as ‘he got’.
- The inchoative auxiliary is lel, E. chudel, as in lilom te kerel, E. chudl’om te kerel ‘I started working’.
- The Western variety has a higher frequency of the passive, as in k’amende terd’i bari khangeri ‘at our place has been built a large church’. Such constructions use intransitive verbs in the Eastern variety: k’amende terd’ol bari khangeri.

2. Second text: A Note on Romani History

a. Words with other alternatives in E: pral (2), E. pal ‘behind, for, about’ (the author uses both forms); čulo, E. čino/čuno/sikra/… ‘little bit’; angoder (5), E. anglunes ‘before’; tham, E. krajos ‘region’ (tham is used in the Eastern area around Svit); bibold’ikan, E. čhind’ikan/židoviko
b. Neologisms
gindo (2) ‘book’ (BD uses alternatively with gendi); chaviben ‘diet’; kheriben ‘living’, BD. avutno ‘origin’ (other authors such as Jan Cibula and Rudolf Dzurko use the word in the meaning ‘future’; this word is derived from avel ‘to come’, hence literally ‘coming’); BD. chalarado (2) ‘specialist’; BD. dżanibmaskero (2) ‘scientific’; BD. phučli ‘question’ (a more common neologism is phučiben); BD. dikhad’i ‘opinion’; BD. nadjanipen (2) ‘ignorance’; BD. mačhengero pañi ‘pond’; šelbersa ‘century’; prindzarinpen ‘piece of knowledge’; BD. prastapengere/prastapune paragrafi ‘deportation laws’.

3. Observations. Two types of words potentially contradict the requirement of “unity” of language. One category is dialectal words. The other is neologisms, which may be different in different dialects and negatively therefore affect communication.

The rate of use of dialectal words can be measured against the total number of words in the texts involved. The total length of text 1 is 1116 words, while text 2 contains 826 words. Out of this number, 29 words (2.6%) for text 1 and the same number (3.5%) for text 2 are affected by dialectal differences at the lexical level. However, in text 2 the number of different words was lower, because some dialectal words (such as mamuj ‘against’ and angoder ‘before’) appeared with considerable frequency. However, comments by native speakers of Romani indicate that the lexical variation does not necessarily render the words concerned incomprehensible.

The second category, neologisms, shows the same percentage of usage in both texts. The number of neologisms in text 1 is 23 (2%), while in text 2 there are 17 newly coined words (2%). This low percentage in the more specialized text 2 was surprising. It seems that a topic such as history does not place requirements on new vocabulary coinage that would be stricter than in the case of communicating about urban daily-life situations. Most of the neologisms are words that are easy to understand. Problems can arise only in the case of words such as avutno, used by Bartoloměj Daniel with the meaning ‘origin’ while some other writers employ it to render the meaning ‘future’.
On the whole, we can say that up to approximately 5.5 percent of the vocabulary of the two texts examined consists of words that are not in the lexicon of speakers of the other dialect of Slovak-and-Czech Romani. Daniel’s lexicon can be further characterized in the following way:

a. He tries to replace loans from Slovak either by original Romani words (archaisms, or words from other dialects, such as E. chudel ‘to get’ for W. dostajinel) or by loans from other languages that do not look Slovak (e.g. rod’ina from Czech for W./E. famel’ia).

b. Although using special Western dialect forms, he also employs common W./E. forms of the words to convey the same meaning. For example, in text 2 he uses W. pral ‘after, about’ on the first two occasions, but following that the word appears consistently in the Eastern form pal.

c. Not unlike other Roma authors, he uses a considerable but not excessive amount of “international” terminology, words that are shared by most European languages: historia ‘history’, etnograf ‘ethnographer’, islamos ‘islam’, etc.; 2.6 percent and 3.6 percent of words in texts 1 and 2 respectively belong to this type.

d. In contrast with many other authors (e.g. Anrej Giňa), Daniel does not translate abundantly from Czech at the word level, and his coinages are easy to accept.

4. Evaluation of the language of the two texts. The existence of linguistic variation is frequently noted by Roma speakers. At the level of stereotypic stories speakers from Levoča are laughed at because they “neigh” (say hi instead of the common ha ‘yes’) and girls from Humenné because they wear to a dance bokanči (in Humenné just ‘shoes’ while in other dialects ‘heavy boots’) (Húbschmannová et al. 1991). Bartoloměj Daniel, the author of the two texts quoted above, is one of the Roma who firmly believe in the correctness of their own variety of Romani. In a review of the Romani-Czech and Czech-Romani Dictionary (originally in the weekly Lidové noviny, reprinted in Amaro lav 8, 4) he criticizes the book for neglecting the Western dialect and says, “Why do we put Slovak words into the Romani-Czech dictionary, if genuine Romani words [for the same concepts] exist?” And further, “An important shortcoming of the dictionary is that it preposes Slovak or Czech prefixes to Romani words.” This is a clear puristic position.

In order to establish reactions to Daniel’s two texts, three Romani speakers of the Eastern dialect were asked to read the text and comment on its language. The first speaker was MS, male, 40 years old, born in Bohemia, whose parents come from Prešov (Eastern Slovakia). RD, who is 43 years old, was born in Bohemia but spent his childhood in a Roma settlement close to Michalovce in Eastern Slovakia. HG is a female
speaker, 23 years old, born in Bohemia but brought up in Moravia, parents originally from the Poprad region (Eastern Slovakia). The last two subjects have passed the Czech matriculation examination. Participants in the interviews, conducted in Romani, also mentioned the language of another author, Ludevit Gábor, whose article was published in the same issue of Romano đžaniben. Unlike Daniel, Gábor is an Eastern dialect speaker.

After praising Daniel for his social stance, MS says,

He wants to show the world that the Roma have, how did he say that, angalalipen [he takes over Daniel’s coinage meaning ‘strong point’] compared with other people in the world. And his language? There are words I didn’t understand. For example, mamuj [‘against’]. What is mamuj? We say prociv. Or the expression oda but’i [‘that work’; a masculine demonstrative with a feminine noun]. As if ‘work’ were a man. That doesn’t fit. Or when I read prastapun paragrafi [‘deportation laws’, one of Daniel’s neologisms], I had to think what he meant. However, if one reads the lot, all of it, it can be figured out what the meaning is. So, his language, his speech is not as beautiful as ours, for example that of Gábor, but it can be understood, all of it can be understood. And if not, one can get used to it. Of course, I have spoken with Western [dialect] people and know how they speak. Vlachi are more difficult to follow but the Western [dialect] people are our Roma.

Notice that although MS claims that he did not understand the Western variety word mamuj ‘against’, he himself gives its correct Eastern dialect equivalent (prociv) in the next sentence. The interview also demonstrates that at least some speakers have developed strategies for dealing with linguistic variation in Romani.

The second interviewee, RD, had the following comment to offer about Daniel:

I know, how he speaks, I have spoken with him several times. When I speak with him, I understand everything, but when I read his stuff, there are words I must think what they mean. They have some words better than we do, but they also have words which are worse than our language. Mamuj [W. ‘against’], that is a nice word, better than prociv [E. ‘against’], which is from Slovak. But čulo [W. ‘a little’]? O gosh. That is a word! This is how Roma who eat dog meat speak. I hate that word ... And Gábor [the other author]? ... His language? The language is that of the Roma from the Košice region. He writes his own words, but I can easily read them.

As noted above, the perception of the word čulo is incorrect: the degeša do not possess a different dialect, the word just belongs to another regional variety of Romani.
HG, when asked whether she understood Daniel’s language, said, “Of course, it is Romani.” When further reminded that it was different from the language she spoke herself, HG elaborated: “It is like Bohemians and Moravians. I was brought up in Moravia, close to Blansko, in a village where non-Roma speak differently from people from Prague. Or in Brno where I studied. Do you know what is ‘šalina’?” The interviewer confirms that she does not know the expression, and HG continues: “You see, it means a tram. So, Daniel, his language is like ours, just a little bit different.” The interviewer: “He says čulo, you say čino.” HG: “That doesn’t matter, does it. You have that in Czech, too: ‘a little bit’, ‘a tiny bit’, ‘a drop’... That doesn’t matter.” HG was further pressed by the interviewer to compare Daniel’s language with the language of the other author, a speaker of the Eastern dialect, but she refused again to evaluate the variation.

The interviews point to some interesting facts. First, while reading, speakers do note individual words. Two have noted the word mamuj ‘against’, but we must consider that it had an especially high frequency in the second text (eight occurrences). It is interesting to see that one of the judges evaluated it negatively, while the other one provided a positive comment. There are other negative evaluations, but they do not affect all noted differences. Only one of the judges identified with the Eastern dialect writer. The youngest of the three speakers, HG, refused attempts to have her make any evaluations. This may be the attitude of the young generation. It should be noted that HG writes and publishes poetry in Romani and her lack of evaluation cannot, therefore, be explained as a consequence of a reduced competence in the language.

On the whole we can say that although we have documented differences as well as the existence of attitudes to variation, there was no indication that such variation would critically impede communication or become the basis of social division that could not be overcome. Of course there is social division, but it does not correlate with linguistic boundaries. To see variation within the Slovak-and-Czech dialect of Romani as in urgent need of standardization would go far beyond the findings of this study.

The spelling. The only feature of Slovak-and-Czech Romani that has been standardized is its spelling. Of course, spelling is in general the component of language the uniformity of which is easiest to achieve (cf. Daneš 1979). However, let us add that such uniformity is often the result of the process of establishing rules that did not exist before. If different spelling rules did exist (as in the case of various Englishes or in the case of Medieval Europe) uniformity would not be so easily achieved.
The Union of Gypsies—Roma, established in Czechoslovakia in 1969, started publishing the journal *Romano l’il*, where Slovak-and-Czech Romani was used for the first time in print. The Linguistic Commission of the Union worked out spelling rules for the variety, based on Slovak and Czech spelling. A slight modification of the spelling rules appeared in *Romano nevo l’il* 88–90 (4–24 October 1993). This is the system used in this paper.

Some of the strategies of the spelling are identical with the spelling recommended by the International Congress of the Roma held in Warsaw in 1990. The International spelling requires more diacritics and special letters than the Slovak-and-Czech variety. The correspondence between the Slovak-and-Czech Romani alphabet and the International alphabet is given in Table 1. The use of the alphabet depends on some morphological criteria, more so in the International than in the Slovak-and-Czech system. In the Slovak-and-Czech spelling mainly the following strategies are applied:

1. Voiced consonants are devoiced at the end of words in pronunciation but not in the spelling.
2. Aspirated consonants also lose their aspiration at the end of words when pronounced but retain the aspiration in spelling.
3. Sometimes elision of vowels takes place and the elided vowel is replaced by an apostrophy (e.g. *andr’amar kher* ‘in our house’).

The most important deviation of the Slovak-and-Czech Romani spelling from the spelling of Slovak and Czech is the absence of the rule that the letter *i* (and sometimes *e*) palatalizes the preceding dental consonant. For example, in Czech *div* is pronounced as if written *d’iv*, and in Slovak *den* is pronounced as if spelled *d’en*. On the other hand, in Romani no dental consonant before *i* (or *e*) is palatalized, unless it carries the “hook.”

Out of the 16 Romani publications launched on the territory of former Czechoslovakia so far 14 use the agreed spelling. Only one of those written in the Slovak-and-Czech variety does not adhere to its rules. Three weekend seminars have been organized so far in the 1990s to discuss matters of spelling and other issues of language.

The adherence to the spelling is basically up to the writers, although sometimes editors correct the writers’ spelling. One of the authors of this paper (Milena Hűbschmannová) edited, for example, Andrej Giňa’s collection of short stories *Bijav* (Prague 1991). Giňa, who writes in the EP variety of Slovak-and-Czech Romani, does not always adhere to the spelling rules. We have analyzed one of his recent stories published in the journal *Romano kurko* (19, 1994). The spelling has not been edited and differs from the system in the following ways:
Table 1. *Letters of the Romani alphabet*

<table>
<thead>
<tr>
<th>Slovak-and-Czech spelling</th>
<th>International spelling*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowels</td>
<td></td>
</tr>
<tr>
<td>a (as in German)</td>
<td>á (Engl. ya)</td>
</tr>
<tr>
<td>e (as in German)</td>
<td>o (Engl. yo)</td>
</tr>
<tr>
<td>i (as in German)</td>
<td></td>
</tr>
<tr>
<td>o (as in German)</td>
<td></td>
</tr>
<tr>
<td>u (as in German)</td>
<td></td>
</tr>
<tr>
<td><strong>Consonants</strong></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
</tr>
<tr>
<td>c (Engl. ts)</td>
<td></td>
</tr>
<tr>
<td>č (Engl. ch)</td>
<td>č</td>
</tr>
<tr>
<td>čh (aspirated č)</td>
<td>ěh</td>
</tr>
<tr>
<td>d</td>
<td></td>
</tr>
<tr>
<td>d' (palatalized d)</td>
<td></td>
</tr>
<tr>
<td>dz</td>
<td></td>
</tr>
<tr>
<td>dž (Engl. j)</td>
<td>ž/dž</td>
</tr>
<tr>
<td>f</td>
<td></td>
</tr>
<tr>
<td>g</td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
</tr>
<tr>
<td>ch (as in German)</td>
<td></td>
</tr>
<tr>
<td>j (Engl. y)</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>k/q</td>
</tr>
<tr>
<td>kh (aspirated k)</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td></td>
</tr>
<tr>
<td>l' (palatalized l)</td>
<td></td>
</tr>
<tr>
<td>m</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td></td>
</tr>
<tr>
<td>ň (palatalized n)</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td></td>
</tr>
<tr>
<td>ph (aspirated p)</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>r/rr</td>
</tr>
<tr>
<td>s</td>
<td></td>
</tr>
<tr>
<td>š (Engl. sh)</td>
<td>š</td>
</tr>
<tr>
<td>t</td>
<td>t/θ</td>
</tr>
<tr>
<td>t' (palatalized t)</td>
<td></td>
</tr>
<tr>
<td>th (aspirated t)</td>
<td></td>
</tr>
<tr>
<td>v</td>
<td></td>
</tr>
<tr>
<td>z</td>
<td></td>
</tr>
<tr>
<td>Ž (as in Engl. genre)</td>
<td></td>
</tr>
</tbody>
</table>

*The international spelling marks palatalization of a consonant also with a wedge placed above the immediately following vowel.*
The spelling system prescribes that voiced consonants that become voiceless in word-final positions are written as voiced, as in dad 'father'. Giňa writes dat.

Similarly, aspirated unvoiced consonants lose their aspiration in word-final positions but this deaspiration is not reflected in spelling, as in jekh 'one'. Giňa writes jek.

The spelling system prescribes that where elision takes place, as much of the original word as possible is retained, as in kija 'to', kij'amende 'to us'. Giňa spells it ki amende.

The negative particles na and ma are written separately from the following verb, as in na kerel 'doesn't do'. Giňa spells it nakerel.

It can be seen that deviations from the spelling system are rather minor. It should be added that Giňa does adhere to some other rules the violation of which is tempting:

- He does not use the grapheme y to denote that the preceding dental is not palatalized. He spells correctly adi 'she' (not ady).
- Although in his variety palatalized d' is pronounced dź, he spells it, in accordance with the recommendations, d': phend'om (not phendžom).
- Similarly, he spells but'i, although he pronounces buči.

It is interesting to note that another article in the same issue of the journal (Romano kurko) by Helena Demeter adheres to the spelling perfectly.

An author who is strongly influenced by Czech spelling is Štefan Červeňák, born in Bohemia but writing in the Eastern dialect. A partial analysis of a collection of proverbs, prepared for the journal Romano gendalos, has revealed a number of deviations from the Slovak-and-Czech Romani spelling: purykane (for phurikane), nič (for ňič), jek maro maškaro deš d'ene (for jekh maro maškar o deš džene), odat edaj amaro d'ivipen (for o dad e daj amaro dživipen), dyk (for dikh), etc. Perhaps a related factor is that Červeňák’s Czech spelling also contains a number of errors.

Overall we can say that apart from examples such as those given above, spelling in journals and other publications has been unified to a remarkable extent. In particular, the use of y has virtually been eliminated and palatalization has systematically been marked with the “hook.” However, deviations appear. Editors do not always correct the spelling. As the editor in chief of Romano kurko has said, “I am the only Rom in the office and there is no time to do all I would like to.”

The standardization of Slovak-and-Czech Romani. The discussion presented above has demonstrated the extent of variation in Romani of the Slovak-and-Czech language group. We have argued that such variation is not extensive and, although noted and evaluated by speakers of the
dialec� concerned, does not seem to impede communication or correlate with social divisiveness. We have also shown that a degree of standardization has been achieved in the spelling of the language, although some variation persists.

These results do not justify a view that further standardization and codification along the lines adhered to in the past by other European languages should be quickly carried out. In section 2 of this paper we reminded the reader that standardization is a process that changes across history. The European processes and the resulting standards belong to the early modern and modern periods of history. At present the situation is changing and new features of standardization and standards have appeared.

We can expect that within postmodern societies standardization will be a process completely different from the modern processes. When considering standardization of Romani, we should not commit ourselves to the old models and should try to apply principles that correspond to the historical context in which the results of the standardization will be used in the future.

In section 2 we quoted several features of the modern concept of standardization. Now these features will be reintroduced and varied to suit the postmodern conditions.

a. Uniformization. The argument for a unified norm is based on modern perceptions of variation, perceptions that evaluate variation negatively and attempt to remove it. An argument that is frequently applied ex post facto to legitimate the concept of a unified norm is the need for interintelligibility. This argument is weak. Examples quoted of dialects that could not be understood by speakers of other dialects are often exaggerated and show that no attempt has been made to overcome the differences through speaker adjustment. It is more likely that the principal driving force of uniformization of language was to support socioeconomic unification of the societies concerned and to destroy a basis on which future separatism might develop.

A long time ago Haugen pointed to the phenomenon of semicommunication (Haugen 1966), a situation in which speakers of two different but cognate languages communicate by using their own language toward speakers of the other language. For example, Danes and Norwegians can communicate in this way. So do Czechs and Slovaks, or Czechs and Poles. Although this way of communication may present problems for those completely unaccustomed to it, after a certain period of adjustment this mode creates no problems at all. In the case of Romani there is no problem of interintelligibility between varieties of the Slovak-and-Czech
group. More adjustment may be needed in the case of other Roma groups such as the Vlachi or Hungarian Romani, but this, in our view, is not impossible to achieve.

Rather than forcibly unifying the language through selecting one of the varieties to serve as the standard, or creating a new standard from individual features of a number of varieties, it is equally possible to train speakers to accept variation and to learn how to deal with it. Our data have shown that native speakers of Romani to a certain extent possess this competence. Problems that speakers perceive when encountering other varieties of the same language mostly result from the lack of such training. Khubchandani (1981) has pointed to Indian communities within which more than ten varieties of language are used without the feeling of inadequacy of any kind. If this degree of tolerance toward variation in the spoken language is possible, there is no reason to assume that a change of attitude (even in the case of undistinguished readers) cannot be achieved in the case of the written language. In any case, there was no strict unification of spelling norms throughout the Middle Ages, and readers, even those who were not highly proficient, had to adjust to variation.

In 1978 Kloss used the term pluricentric language to describe a situation in which several national varieties of the same language, each with its own standard norms, interact (Clyne 1992). The term could also be used to describe the situation of a number of dialects in which intercommunication is practiced. We believe that this is the model to be used when considering the issue of standardizing Romani.

b. Elaboration. Modern standards are not only unified, they are also elaborated. The lack of elaboration is noted by contemporary speakers of Romani, is negatively evaluated, and adjustment processes initiated. The most typical of these processes is the process of word formation, creation of new terminologies. Evidence for the need of such processes in Romani has been revealed above in our discussion of neologisms in the two texts studied. These processes exist as natural-language management processes that develop on the basis of actual needs. Unlike the issue of uniformization, elaboration will remain a real issue for Romani.

At present the situations in which the need for elaboration (or development) of Romani is most keenly perceived are situations in which writers and journalists are placed. As soon as Romani is introduced into schools, the development of the school variety will be needed. Let us not forget that although there are many families in urban settings where the language is not any more the language of primary socialization, in some districts of Eastern Slovakia children still arrive at school with Romani as their
only language. At present Romani is not used in schools, either in the Czech or in the Slovak Republic, except for occasional reading of Romani stories in special classes for Roma children.

Radio and television are another area in which language development will be needed. Both Czech and Slovak television once a week broadcast the program Romale. This is a 30-minute program, timetabled after the midnight news and produced entirely in Czech and Slovak. Roma activists would like to have Romani used at least for a part of the program but are afraid to openly suggest this because the whole program might be scrapped.

Development of the language is also needed for the social life of the Roma, national and international conferences, and meetings of the International Romani Union, an organization that holds congresses and conferences and publishes a newsletter partly written in Romani.

If elaboration of the dialects is to be taken seriously and it aims at retaining flexibility as well as adjusting to local conditions, full uniformity across national boundaries cannot be expected; understandably, an attempt should be made to avoid diversification where it can be avoided. However, we cannot wait for the development of an international standard of Romani lest we risk that, in the meantime, the language disappears. The international standard, if needed, can be developed alongside the pluricentric standard (cf. Clyne 1992).

The elaboration of Romani should not set for itself aims that greatly surpass current needs. For example, we can see no need, at least at this stage, for complicated scientific terminology. Romani is functionally complementary with languages of the matrix communities and it would be unrealistic to suggest at present that this will not be true in the near future. In other words, elaboration of standards in the postmodern period may be selective. To claim that Romani should be elaborated as a full national language might lead to phenomena we know from some postwar developing languages, in which the process of elaboration preceded the development of needs.

c. Codification. The process of elaboration should be recorded and its results should be widely available. However, this does not imply codification. The recorded rules do not have to be followed. They are for reference only.

d. Base in contemporary spoken language. Modern standards are based on the spoken language of the whole or part of the population. We can assume that the principle of postmodern standardization will remain the same. This has two particular aspects.
First, we can expect a certain amount of purism, such as manifest in the case of Bartoloměj Daniel quoted above. Purism is not necessarily bad (Neustupný 1989): what should be watched is some social sources of purism, such as the aim of separating oneself from the matrix communities. Here the actual situation of the communities in question must be considered. However, it would be both unrealistic and contradictory to the interests of the Roma masses to require that all elements resulting from the influence of the matrix languages should be purged.

Second, the principle implies that standardization must take account of the matrix languages. A certain degree of mixing is unavoidable and should not be faced with fear. At the same time we must realize that this principle reinforces the trend to further diversify individual dialects of Romani spoken in areas with a different matrix language (Czech, Slovak, Hungarian, Rumanian, etc.).

e. Symbolic values. Modern standards served as symbols of the unity of national economies, societies, and cultures. Within the context of postmodern international integration, it would be unrealistic to posit the need for the economic and political symbolism of Romani. It can be expected that the basic symbolic function will be a cultural one: a symbol of the Roma contribution to the world culture. This purpose can be served by individual dialects, or the sum of the dialects, without developing one of them into the old-type standard.

The first group of the Roma for whom this symbolic value is of interest is the newly developing middle class. None of the numerous Roma organizations registered in the territory of former Czechoslovakia leaves the requirement of the use and development of the language out of its programmatic statements. Of course, such statements are normally authored by those Roma who aspire to a middle-class status. In practice, linguistic aims do not figure prominently in the activities of the organizations.

It will be necessary to watch the situation closely. On one hand, it is impossible to eliminate the symbolic function of language. On the other hand, the symbolic function can be performed without any real competence in the language. In other words, language can serve as a symbol even if it is not actually used and if no commitment exists to its development.

f. Whose property is the standard? It has often been pointed out that modern standards are in fact the property of the middle class (cf. Tollefson 1991), which then makes them available on loan to the lower classes of the society.
Written Romani is no exception. Compare MS’s attitude to Daniel’s language with that of the other two judges. MS is a laborer, the other two are, by Roma standards, the middle class.

In a postmodern society the matter of ownership is important. It may be expected that the distribution will be quite wide, and those who work on the Romani standard should pay due attention to the issue. Is the newly created lexicon available to all or only to those who regularly read and act as gate-watchers of the written language? The other features of standardization discussed above are related to this issue; for example, the way the language is elaborated, its relationship with the matrix variety, or its symbolic function.

Conclusion

Old-fashioned language-management executives may push the standardization of Romani into the old mold, which is familiar to them. However, the modern model of standardization should not be allowed to come back. The new model of standardization expounded in this paper can be summarized as follows:

1. The standard is polycentric.
2. The standard is selectively elaborated.
3. There is no codification that would be binding for participants.
4. The standard is a mixed home language.
5. The standard is a symbol of ethnic contribution to the world.
6. The standard is a property of all, not only the elites.

Charles University, Prague
Osaka University

Notes

1. The term Slovak-and-Czech dialect will be explained later in the text.
2. The term matrix is used here in the sense it has, for example, in geology (the overall mass encompassing other elements). If there is any value judgment, it implies guilt through dominance.
3. This paper spells Romani in accordance with the Slovak-and-Czech Romani spelling system. For more details see the comparative table of the Romani alphabet given later in the text.

References


Havránek, B. (1963 [1932]). Úkoly spisovného jazyka a jeho kultura. In *Studie o spisovném jazyce*, B. Havránek, 30–59. Prague: Academia. (Section 2 has been translated into English in Garvin 1964.)


