Anna-Maria Meyer

The creation of orthographies for Romani
by means of 'Slavic' alphabets

Abstract

Romani, the language of the Sinti and Roma, is spoken all over Central, Eastern and South Eastern Europe, i.e. also in numerous countries with a Slavic-speaking majority. Although there have been various attempts of finding a standardized script and orthography for Romani worldwide, there is now a clear tendency away from international towards national or regional projects of creating alphabets and orthographies for Romani. In the present paper, ten such attempts from seven different Slavic-speaking countries are analyzed and compared against the background of Smalley's (1964) and Coulmas' (1989) theoretical considerations about the creation of alphabets for hitherto unwritten languages. The questions to be answered are: Which script and which orthography serve as examples for each proposal? Which solutions are found to represent the phonetic features of Romani by means of Slavic alphabets? And: Have the proposals been accepted by the speech community?

Keywords: Romani; Slavic; Latin script; Cyrillic script; orthography; standardization

1. Introduction

Traditionally, Romani (or Romanes) has been passed on through the generations exclusively orally, its written form began to play a larger role only in the course of the 20th century. Until about a hundred years ago, no written records by Roma themselves were conveyed. The earliest attempts to find a consistent orthography for Romani are from non-Roma (*gaje* in Romani), who wrote down the language by ear and against the background of their own language. The first proof of written Romani dates back to the year 1547 in England (cf. Bartosz 2009: 154). However, in the past few decades, especially since the political turnover in 1990, the Roma have used their language in a written form more and more naturally, and nowadays writing is by far no exception for Roma any more (cf. Matras 1999b: 97).

In spite of different attempts no worldwide standard for a written Romani has asserted itself so far. Instead, Romani was and is very often written down spontaneously, which has led to innumerable different spellings; Bartosz (2004: 115), citing the Romani poet Rajko Djurić, enumerates a dozen spelling varieties for the collocation *romani čhib*, as the Roma themselves call their language: *romani czib, romani cib, romani tschib, romani tschiw, romani tsiw, romani tsiv, romani tscheeb, rromani chib, romani sib, romani sip, xomani ćip, romani chib, rhomani ćhib, romanyj sip, kxomani tchib*. Versions in other scripts than Latin are not even included in this enumeration. This great variety along with the lack of a state territory, a central government and sufficient financial and organizational resources are disadvantageous for the widespread dissemination of a standard Romani orthography. However, there are many smaller, regional or national projects that have been developed by small, local task forces and more or less successfully applied in practice.

The subject matter of this contribution are the proposals for creating an alphabet for Romani in countries with a Slavic-speaking majority, an area where many Romani people have been living for centuries. First, I am going to give an introduction to the (mainly, but not exclusively) oral culture of the Romani people and their general attitude to writing. Thereupon I am going to present the ideas of Smalley (1964) and Coulmas (1989) on the creation of alphabets for hitherto unwritten languages, which constitute the theoretical basis for the later analysis. It will also be worked out which of the criteria named by Smalley and Coulmas are especially important for Romani in Slavic-speaking countries, where either the Latin or the Cyrillic script – or both – with different orthographies are in use. On this theoretical basis I am going to present and compare ten proposals for the creation of an alphabet for Romani from seven different Slavic countries: Russia, Ukraine, Czech Republic, Slovakia, Macedonia, Bulgaria and Poland. The following questions should be answered: which script and which orthography serve as examples for each proposal? Which solutions are found to represent the phonetic features of Romani by means of Slavic scripts and orthographies? And: Have the proposals been accepted by the speech community?

2. The role of literacy in Romani culture

Talking about written Romani, it is, first of all, important to understand the role of orality and literacy in Romani culture. Referring to Ong (1982), it has to be greatly emphasized that primarily oral cultures should not be seen as something primitive or deficient, but as equal to literal cultures. The problem of a – conscious or unconscious – depreciation of orality can simply be explained by the difficulties of a literal society to imagine how an oral culture works, which leads to sometimes helpless comparisons like explaining the meaning of "horse" departing from the idea of a car (cf. ibid.: 12f.). The sound in an oral culture has a different relationship to time, it exists only for a moment and cannot be conserved, so the spoken word is of great weight and importance, it is a guide to social action. Ong (ibid.: 32) refers to the Hebrew word *dabar* in this context, meaning both 'word' and 'event'.

2.1. Romani culture as an oral culture?

In Romani culture, often narrative and lyric forms (fairy tales, riddles, proverbs, songs etc.) are used to orally convey important information which regulates the life in the group and the part every member plays in it. Thus, the cultural identity is sharpened and passed on from generation to generation (cf. Toninato 2014: 46; 48). For a very long time, Romani culture has been perceived as an exclusively oral culture – "nicht zuletzt deshalb, weil sich diese Zuschreibung [...] problemlos mit den essenzialistischen Meinungen über kulturelle Rückständigkeit und Anpassungsfähigkeit verbinden liess [sic]"[1] (Kurth 2008: 77). Toninato, as well as Kurth, suggests that one should not strictly separate orality and literacy, because of the countless hybrid forms:

> [T]he lack of an alphabetical writing system does not automatically imply the absence of all forms of writing, since writing is a much broader phenomenon encompassing the production and the use of graphic systems for communicative purposes (Toninato 2014: 53).

---

[1] 'not least because this ascription [...] could be linked smoothly with the essentialistic opinions about cultural backwardness and adaptability.'

It would be correct to say: communication that is based on alphabetic writing is not central to Romani culture; however, the Roma have developed their own ways of passing on information in other than oral ways over the centuries, above all non-alphabetic graphic signs.

A historical example can be found in Piasere (1985: 162): when a group of Roma was travelling and on its way had nothing extraordinary to report to the following group, this was signalled by tying three bunches of herbs to a branch by the wayside or crossroads, with a small stone tied to the first bunch. In dangerous situations, for example on escape, the following groups were warned by scattering the herbs on the ground in the middle of the road. Among the Slovene Roma, this way of communication is called *tragi* 'signs', often also *patrin* (the Romani word for 'leaf', cf. Toninato 2014: 55–59). The linguist and ethnographer Heinrich von Wlislocki (1994: 152) wrote down his observations about a "group of Upper Hungarian travelling gypsies", whose leader wanted to communicate to the following group(s) the place where he had stayed on Wednesday after the fifth Sunday after Whitsun: a rag is tied to a tree in the direction of travelling; it is provided with five (for the number of Sundays) stitches of red wool (the sign of the leader) lengthwise and three stitches across (for the three days of the week). Romani women at Wlislocki's times additionally used so-called *čine* signs (*čine* is the name for vagrant activities of women such as selling small items, begging and fortune-telling), that were written on house walls. Wlislocki (ibid.: 144) mentions, among others, a double cross, which means "inhumane treatment", or a double circle, which means "very good people". Both kinds of signs are closely linked with the traditional vagrant lifestyle of the Roma and are invisible for *gaje*. With the increasing sedentary lifestyle, these forms of communication slowly but surely have lost their importance.

Furthermore, the Roma have at all times been under the influence of the literal cultures surrounding them, even though the majority of them was not literate themselves (cf. Toninato 2014: 1). If it was necessary for them, they were able to write; however, for a long time they deliberately avoided taking over the literacy of the majority population in order to keep away foreign influence from their own culture. Only a few decades ago the approach of the Roma to writing started to change. Consequently

they are facing the dilemma of wanting to keep their own, valued culture but being less and less able to prevent themselves from omnipresent literacy, which also has to offer better chances of education and employment. The critical attitude of many Roma to schooling guarantees the preservation of their culture to a certain degree, but on the other hand hinders their chances of social promotion, which brings with it economic independence and more integration into the society. It is important to understand that the transition from orality to literacy does not simply mean a change in the way of communication, but it brings with it profound changes in the social structures. Against this background it is difficult to decide whether literacy is rather useful or harmful for the Roma. Ong (1982: 15) sums up the dilemma as follows:

> This awareness is agony for persons rooted in primary orality, who want literacy passionately but who also know very well that moving into the exciting world of literacy means leaving behind much that is exciting and deeply loved in the earlier oral world. We have to die to continue living.

2.2. Possibilities of writing Romani

In spite of the just depicted dilemma, the writing of Romani seems to be inexorable and has in fact been practiced for many decades already. Apart from letters, poetic and narrative anthologies, texts with symbolic function such as bible translations and political publications,[2] especially online communication in Romani is gaining ground:

> The establishment of Romani-language websites and email discussion lists from around 1995 onwards has changed the face of written communication in Romani completely. It is impossible to estimate the number of Romani-language email users; the figure is definitely rising rapidly. Email has given Rom from different countries, who do not necessarily share a second language, a medium for spontaneous written communication in Romani (Matras 2002: 257).

Very often and especially on the internet, Romani is scripted spontaneously by way of the alphabet and orthography that is known to the writer, i.e. the alphabet and orthography of the majority language. This is the

---

[2] On text forms and motivations for writing in Romani cf. Matras (1999b).

most unsystematic, but also the most widespread solution to write the language.

To find a more systematic approach, it has been attempted since the 1970s to standardize Romani. On the IV International Romani Congress in Serock near Warsaw in 1990, a proposal for a worldwide uniform orthography of Romani was submitted by the language commission of the International Romani Union (IRU)[3] on the basis of the Latin alphabet. In charge was Marcel Cortiade (Courthiade). This alphabet, consisting of the graphemes <a, b, c, ć, ćh, d, e, f, g, h, x, i, j, k, kh, l, m, n, o, p, r, rr, s, ś, t, th, u, v, z, θ, ʒ>, along with a system of orthographic rules, was officially accepted by a congregation of Roma from 20 countries on April 8th, 1990. The project and its history are elaborately described in Courthiade (2012) and elsewhere. Igla (1991: 87) criticizes especially the "abstract morphographemes" of the Warsaw alphabet; see also the criticism in Kochanowski (1995). It has been used in practice,[4] but not as widespread as it was hoped by its author. Less well-known, but worth mentioning is the attempt for standardization by the Spanish Gitano Juan de Dios Ramirez Heredia, which has, however, not been picked up by other Roma (cf. Matras 1997: 112). Apart from this proposal, also Kenrick (1981) and others have been discussing possibilities of finding a standardized orthography for the international Romani community, however, a generally accepted solution has not been found so far.

The only area in which a relatively uniform standard of Romani could be established is linguistics. Although it has never been codified, a convention for writing Romani in linguistic works has developed over the decades. It makes use of the Latin alphabet, háčeks[5] for the notation of palato-alveolar consonants (<š, ž, č>), the grapheme <h> for aspirated

---

[3] The IRU (in Romani: *Internacionalno Jekhetanibe*) was founded in 1978 during the II International Romani Congress in Geneva and is the most important international union of Romani people.

[4] E.g. in Romania, where a government-supported initiative introduced the alphabet for the school curriculum in dozens of schools with a high number of Romani children (cf. Matras 2014: 125). Furthermore, the IRU as well as the academic journal *Studia Romologica* (www.studiaromologica.pl/wp-content/uploads/studia_romologica-2-2009.pdf) use it for international publications.

[5] Instead of the word *háček*, which is common in Slavic philology, in romological literature the term *čiriklo* is in use, which means 'little bird' in Romani.

consonants (<kh, ph, th, čh>) and <x> for the velar fricative. However, even here we find deviations from the established convention (cf. Matras 1999a: 488; 2002: 254; Heinschink & Cech 2013: 72).

There is an ongoing trend that leads away from international standardization of Romani orthography towards smaller, national or regional approaches. The examples from the present paper underline this development. Hübschmannová & Neustupný formulated already in 1996 rather drastically but correctly:

> [...], 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 (Hübschmannová & Neustupný 1996: 105).

In accordance with this "pluricentric standard" (Friedman 2005: 163 calls it "polycentric"), writers of Romani often use the alphabet and orthography of the majority language in their country. Matras (1997: 114) takes it as an advantage that these projects are smaller in size, involve a manageable circle of people who support the idea and are ready to actively work on it. Although there are usually one or two people in charge, these projects are not one-man endeavors, which is beneficial for their success. Furthermore, they address the speech community in their own dialect and with an alphabet that can be learned without much effort due to its proximity to that of the majority language of the country.

3. Scripts and orthographies for Romani in seven Slavic-speaking countries

By the somewhat simplified formulation "creating an alphabet" I understand the introduction of a script and an orthography for a hitherto unwritten language, which in German is termed *Verschriftung* (cf. Buß- mann 2008: 667). This scripting can stand for itself or can be part of a larger standardization process which also involves other levels of the langauge. Challenges and possibilities that can appear during this process are presented well-foundedly in Smalley (1964) and, based on that, Coulmas (1989: 225–240). Hence, these two authors shall serve as the theoretical basis of the later analysis.

3.1. Theoretical background

To Coulmas (1989: 226), a good orthography is more than a mere transcription system. Writing systems and orthographies are emotionally charged, they represent a mirror of the identity of the speech community in question and possess strong symbolic power – therefore, not only linguistic, but also social aspects have to be considered. Smalley (1964) compiles a list of five criteria which is discussed and further elaborated by Coulmas:

1. Maximum motivation for the learner. By this, Smalley and Coulmas refer to the above mentioned fact that the speech community has to accept the writing system, i.e. that its success does not depend on linguistic factors only. Language attitudes play an important role, especially when it comes to the imitation of or – on the contrary – demarcation from another writing system, usually that of the majority population of the country in question. It can either serve as a prestigeous model worth copying or as a negative example to deviate from (cf. Coulmas 1989: 227).

2. Maximum representation of speech. The idea behind this criterium is that an orthography should represent the spoken language as faithfully as possible, which, however, is not easy to apply in practice. First, the problem of language-internal variation must be solved, i.e. the question must be answered which dialect should serve as the basis. This alone can lead to many difficulties through potential discrimination of a part of the speech community (cf. Coulmas 1989: 229). Romani splits into a huge amount of varieties worldwide, whereas no dialect possesses an especially high supra-regional prestige and would therefore be more appropriate than others. Once this question is solved, the criterium of maximum representation of spoken language can, according to Smalley (1964), be achieved best by a phonemic transcription, i.e. by the principle that every phoneme is represented by exactly one grapheme. Coulmas sees no reason to refuse this approach in general, however, he indicates that orthographies empirically move further away from simple phonemic representation over the years:

> [M]ature alphabetic orthographies encode morphological and lexical information in addition to phonemic information; and mature readers make

use of this information more than they do of letter-sound correspondences (Coulmas 1989: 230).

Therefore he – at least theoretically – pleads not only for a phonological, but also for a thorough morphological, syntactical and lexical analysis of the respective language. Practically, he is aware of the complexity of this endeavor, wherefore a focus on phonology is justifiable to him. However, this seemingly simple solution also bears certain challenges: How, for example, should phonemes be represented, for which there are no equivalents in the source alphabet? For the Latin alphabet, different solutions have been found in the past, among them the adoption of signs from the International Phonetic Alphabet, letter combinations, the reinterpretation of 'superfluous' letters, letters in different sizes, sub- or superscript, diacritic signs or different fonts (cf. Coulmas 1989: 231).

3. Maximum ease of learning. This seemingly trivial postulate that an orthography should be easy to learn, at first glance speaks for a phonemic transcription. However, it is not without difficulty to segment the phonemes correctly and then assign a grapheme to every phoneme; some segments are impossible to identify as separate sounds. Besides, it has to be considered that writing and reading make different demands. What eases reading can impede writing and vice versa. Accordingly, for speakers of a language with a distinctively morphological-etymological diction such as Polish, the different spelling of phonetically identical words such as <może> 'maybe' and <morze> 'sea' (both pronounced [ˈmɔʒɛ]) can be helpful in reading comprehension, but a frequent source of errors in writing. Trying to find a compromise, Coulmas (1989: 233) eventually recommends adaption to the readers rather than to the writers – which would in turn speak against a solely phonemic diction.

4. Maximum transfer. Generally, languages that are scripted for the first time in modern times are unlikely to be widely used in communication. As a rule, the wish to script it is nevertheless based on the rational idea that speakers, once they will be alphabetized in their mother tongue, will also be more easily alphabetized in other languages. This hope can, however, only be fulfilled when alphabet and orthography of the language in question show great similarity with the surrounding majority language. When it comes to the Roma, the situation is reverse: They are usually alphabetized in the majority language first and only la-

ter, if at all, in Romani, whereby the orthography of the majority language has an intermediary function. Direct transfer in orthography can be warranted most easily if the graphemes of the hitherto unwritten language represent the same phonemes as those in the majority language. Vice versa, no graphemes from the majority language should be used for which there are no counterparts in the phonology of the hitherto unwritten language. For phonemes without a counterpart in the alphabet of the majority language, new letters or letter combinations have to be introduced (see above). Hence the alphabet and orthography of the surrounding majority language can serve as a basis which can be adapted if necessary (cf. Coulmas 1989: 235).

In the case at hand that means: when Romani is compared to its Slavic contact languages, the following special features of the language have to be considered: Like the Slavic languages, Romani possesses the vowels /a, e, i, o, u/, which are inherited from Indo-Aryan and can be found in all dialects. Further phonemes that exist in only some varieties such as the schwa in some dialects, e.g. in Bulgaria, are contact-induced. The same applies to vowel length. The diphtongs /aj, oj, ej/ do exist, but can only be found in very few inherited words. The most important peculiarity of the consonant system is the aspiration of /p, t, k, tʃ/ > /pʰ, tʰ, kʰ, tʃʰ/), an inherited feature of Romani that has been preserved in almost every dialect, e.g. *phral* 'brother', *thud* 'milk', *kher* 'house', *čhaj* 'daughter'. This feature clearly identifies Romani as an Indo-Aryan language and poses a certain challenge to anyone trying to script it. The phoneme /ʒ/ is rather marginal and limited to loan words. Many, but not all dialects palatalize consonants before front vowels and some possess palatal phonemes such as /t͡ɕ/ and /d͡ʑ/. Some conservative dialects in South-Eastern Europe have preserved two rhotic consonants: a 'rolled' /r/ and a further /r/-sound that can be pronounced [ʀ], [ɣ] or retroflex [ɽ]. Early Romani had a sonorant /ɽ/, which traces back to historical retroflex consonants, and in Bulgaria's Rhodope region such a retroflex pronunciation exists until today (cf. Igla 1997: 152). Elsewhere, this sound has merged with the 'rolled' [r] that is typical of most Slavic languages as well, so that there is no phonemic distinction between them any more. A further challenge is the differentiation of the closely

related phonemes /h/ and /x/ as well as the question whether to represent final devoicing in writing or not.

5. Maximum ease of reproduction. This last criterium is a purely technical one, but it still plays a decisive role in the process of scripting an unwritten language. The graphemes must be easy to write on a computer without complicated additional characters that often cannot be realized by Romani writers on the devices available to them.

Smalley's five criteria are thus to a certain extent in conflict with each other and cannot all be equally respected. Above this, much depends on the language in question, in our case Romani. However, Coulmas emphasizes that the prestige and the acceptance of a diction is the decisive factor in the end: "Linguistic analysis can be of great service and should be the foundation of any new orthography but it can only serve, it cannot dominate" (Coulmas 1989: 238).

3.2. Projects of scripting Romani in slavophone countries

On the basis of Smalley (1964) and Coulmas (1989) we are now going to analyze and compare the national and regional attempts of scripting Romani in the Slavic-speaking countries. Some projects are penned by Roma, some by *gaje*, some by mixed groups. Some of the projects have already been subjects of linguistic research before, some (especially those younger than 20 years) have not; in any case a comparative approach in this constellation and against the mentioned theoretical background is a novelty and can gain interesting new insights. The aim of the present paper is to revive the discussion about the standardization of Romani and to add some new facts and ideas to it.

3.2.1. Russia, Ukraine

In Russia, publications in Romani were released surprisingly early, and after the foundation of the Soviet Union a liberal nationality policy with the aim of loyalty to and identification with the new state was introduced (cf. Matras 2014: 123). After the October Revolution Roma in the Soviet Union had the possibility to attend school and cultivate their language. When in 1927 the first newspaper in Romani (*Zora*) appeared and several schools were opened where Romani children were taught by Romani teachers, it became necessary to script the language and publish books

and teaching material in it (cf. Djurić 2002: 33). From the early 1920s onwards, translations of Russian classics, of the bible, children's literature as well as political pamphlets were published in Romani. In the course of this development, in 1938 North Russian Romani was scripted as the first Romani dialect in Eastern Europe by Sergievskij & Barannikov,[6] who used the Cyrillic alphabet. Some, however very scarce notes on finding an alphabet for North Russian Romani can already be found in Sergievskij (1931: 9). The authors are aware of their pioneering work and broach the issue in the preface of their *Cygansko-russkij slovar'*:

> Цыганский язык до революции был совершенно бесписьменным (каким он остается и сейчас в капиталистических странах). В нашей стране благодаря ленинско-сталинской национальной политике советской власти цыганы имеют свою письменность и литературный язык (Sergievskij & Barannikov 1938: 3).[7]

The authors choose the Russian Cyrillic alphabet as the basis that is slightly adapted to the needs of North Russian Romani. The alphabet reads: <Аа, Бб, Вв, Гг, Ѓг, *Дд*, Ее, Ёё, Жж, Зз, Ии, Йй, Кк, *Лл*, Мм, Нн, Оо, Пп, Рр, Сс, Тт, Уу, Фф, Хх, Цц, Чч, Шш, ы, ь, Ээ, Жж, Яя> (cf. ibid.: 153). Compared to the Russian alphabet, <Ъъ> and <Щщ> are missing, newly introduced is the grapheme <Ѓг>, which exists in the Ukrainian, but not in the Russian Cyrillic alphabet. It stands for the voiced velar stop /g/ as in <гирил> [gi'ril] 'peas'. The aspirated consonants are written by means of the digraphs <кх, пх, тх>, e.g. <пхаро́> [pʰa'ro] 'heavy', <тхуд> [tʰut] 'milk', <кхэр> [kʰɛr] 'house' (cf. ibid.: 154). The example <тхуд> [tʰut] also shows that final devoicing is not represented in writing. If the palatalization of a consonant does not happen 'automatically' through the influence of the following front or palatalized vowel (<е, и, ё, ю, я>), a soft sign has to be inserted as in <бельве́ль> [bʲelʲiˈvʲelʲi] 'Wind'. If the /ʃʲʃʲ/ sound does appear in a loan

---

word from Russian, <шш> is written as in <ба́ршшё> ['baršj:jo] 'borscht' (the example also shows: <ё> is not necessarily stressed as in Russian) (cf. ibid.: 155).

Sergievskij & Barannikov's proposal later served as an example for other authors over many years, which hints at its acceptance in at least parts of the speech community. Among others, Machotin (1993: 5) refers to it in his dictionary and merely reintroduces <Ъъ> and <Щщ>. Šapoval does the same, giving the following explanation:

> Надо сказать, что эти искусственные ограничения были отчасти воплощением теоретических принципов создателей алфавита, а отчасти были вызваны орфографической модой текущего момента (неприятие буквы Ъ в послереволюционной России). Впоследствии они не закрепились, и российские цыгане при записи своей речи на практике не отказались от букв Щ и Ъ (Šapoval 2007: 15).[8]

Through the change in the nationality policy under Stalin the positive development for the Roma in the Soviet Union came to an end. Publications in Romani were forbidden and the attempts to script and standardize it paused for many years. Only in the late 1960s the first new initiatives came to life (cf. Matras 2014: 124; Toninato 2014: 76). In 1990, Demeter & Demeter published their Russian–Romani dictionary for the dialect of the Kalderash.[9] According to the authors, the idea of creating an alphabet for this dialect had been existing since the 1950s, but could not be published until 1990 due to the political circumstances (cf. Demeter & Demeter 1990: 8). The authors appreciate the work of their forerunners Sergievskij & Barannikov and rely on it in many respects, they only replace <Гг> by <Ғғ>. By doing so, they demonstrate a closer proximity to Turkic languages spoken in Russia such as Bashkir and Kazakh rather

---

[8] 'It must be said that these artificial limitations were partially the embodiment of the authors' theoretical principles, and partially they were evoked by the present orthographic fashion (the non-adoption of the letter Ъ in post-revolutionary Russia). Consequently, they have not stabilized and the Russian gypsies in practice have not resigned from using the letters Щ and Ъ.'

[9] R.S. Demeter (1920–1989) is a Romani poet, folklorist and ethnographer with a Ph.D. in pedagogy. His brother, P.S. Demeter, is a composer. They come from a mixed Kalderash-Servika Roma family (cf. https://ru.wikipedia.org/wiki/Деметер,_Роман_Степанович). – The Kalderash are a group of Roma from Romania and South Eastern Europe. Their name derives from the Romanian word *căldărar* 'kettle' and refers to their traditional profession: coppersmiths and boilermakers.

than to Ukrainian. Furthermore, they consider two different realizations of /r/ in their diction: <p> and <pp>. Demeter & Demeter explicitly aim at a possibly close proximity to the Russian Cyrillic orthography and put special emphasis on the phonemic principle: „каждой фонеме (фонетическому явлению) – свой знак, с тем, чтобы один и тот же знак не служил для обозначения разных фонем"[10] (ibid.: 9).

Цыганский алфавит	Страница	В международной транскрипции соответствует	Цыганский алфавит	Страница	В международной транскрипции соответствует
Gypsy alphabet	Page	International transcription	Gypsy alphabet	Page	International transcription
А а	21	A a	П п	114	P p
Б б	31	B b	ПХ пх	128	PH ph
В в	42	V v	Р р	132	R r
Г г	51	G g	РР рр	135	RR rr
Ғ ғ	57	H h	С с	137	S s
Д д	60	D d	Т т	148	T t
Е е	—	ye	ТХ тх	155	TH th
Ё ё	—	yo	У у	156	U u
Ж ж	70	Ž ž, Ź ź	Ф ф	157	F f
З з	72	Z z	Х х	160	X x
И и	74	I i	Ц ц	164	C c
Й й	76	Y y	Ч ч	166	Č č
К к	78	K k	Ш ш	171	S š Ś ś
КХ кх	91	KH kh	Ы ы	178	Ï ï
Л л	93	L l	Э э	179	Ə ə
М м	98	M m	Ю ю	179	yu
Н н	108	N n	Я я	179	ya
О о	112	O o			

Fig. 1: Demeter & Demeter's (1990, n.pag.) alphabet in the Cyrillic original and the Latin transliteration

[10] 'To every phoneme (phonetic phenomenon) one sign, so that one and the same sign will not serve the purpose of marking different phonemes.'

Международная транскрипция	Страница	Соответствие в цыганском алфавите	Международная транскрипция	Страница	Соответствие в цыганском алфавите
International transcription	Page	Gypsy alphabet	International transcription	Page	Gypsy alphabet
A a	233	А а	O o	259	О о
B в	235	Б б	P p	260	П п
C c	238	Ц ц	PH ph	264	ПХ пх
Č č	239	Ч ч	R r	265	Р р
D d	241	Д д	RR rr	266	РР рр
E e	—	Е е	S s	267	С с
F f	243	Ф ф	Š š, Ś ś	271, 272	Ш ш
G g	244	Г г	T t	273	Т т
H h	246	F ғ	TH th	275	ТХ тх
I i	247	И и	U u	275	У у
Ï ï	248	Ы ы	V v	276	В в
K k	248	К к	X x	278	Х х
KH kh	252	КХ кх	Y y	279	Й й
L l	253	Л л	Z z	279	З з
M m	255	М м	Ž ž, Ź ź	280	Ж ж
N n	258	Н н	Ә ә	281	Э э

Fig. 1 cont'd: Demeter & Demeter's (1990) alphabet in the Cyrillic original and the Latin transliteration

Yotized vowels can, however (which contradicts this principle), be represented in two ways: either as <я, е, ё, ю> or as <йа, йе, йо, йу>. By request of international researchers the authors present all entries also in Latin transliteration (see fig. 1). A decade later, Cvetkov (2001: 20) refers to Demeter & Demeter in his dictionary for the dialect of the Rus-

sian Lovari.[11] He only adapts the diction in so far as the letter <Г'г'> is introduced to mark /h/ and double vowels are introduced to mark vowel length, as in <пативаало> 'upright, honest'.

The latest project of creating an alphabet for Romani in the East Slavic area was published by Toropov & Gumeroglyj[12] (2013) and refers to the dialect of the Crimean Roma. A basic principle for the authors is the exclusive use of graphemes from the modern Russian-Cyrillic alphabet, at the same time they emphasize that their diction is something of its own:

> Каждый человек, пишущий на языке крымских цыган, должен понять, что он пишет на цыганском языке, используя оригинальный цыганский алфавит, составленный из кириллических букв, со специфическими, только ему присущими, правилами алфавита и орфографии, а не записывает цыганские слова по правилам орфографии русского языка, как это *делали некоторые российские лингвисты до 1918 г. Например, цыганские слова пхол* ,золото' и *кхам* ,солнце' в своей публикации от 1875 г., этнограф В. Х. Кондараки записал следующим образом: *холъ* ,золото' и *камъ* ,солнце' […]. Буква *ъ* в цыганских словах в этом конкретном случае – дань русской орфографии того времени, а не обозначение какого-либо цыганского звука (Toropov & Gumeroglyj 2013: 202).[13]

Abiding by this principle without additional diacritic signs facilitates writing Romani on a computer (cf. ibid.: 199). The authors suggest a combination of the phonematic and the etymological principle and emphasize the advantage of the smaller grapheme inventory in the former. Final devoicing is not represented in writing, e.g. [gat] 'shirt' should be

---

[11] The Lovari are a group of Roma closely related to the above named Kalderash. Their name derives from the Hungarian word *ló* 'horse' because of their specialization in horse trading.

[12] P.B. Gumeroglyj (1960–1999) was a Crimean Rom. He and the Russian V.G. Toropov got to know each other in 1979 and discovered their common interest in the language and culture of the Crimean Roma (cf. Toropov 2003: 3).

[13] 'Everybody who writes the language of the Crimean gypsies should understand that he is writing in the language of the gypsies and using the original gypsy alphabet, consisting of Cyrillic letters with specific, typical rules of the alphabet and orthography, and that he does not write down gypsy words according to the rules of the Russian language as was done by some Russian linguists until 1918. For example, the gypsy words *phol* 'gold' and *kham* 'sun' were written down by the ethnographer V. Ch. Kondaraki in his publication from 1875 as follows: *холъ* 'gold' and *камъ* 'sun' [...]. The letter *ъ* in gypsy words is in this case a tribute to the Russian orthography of that time and not the notation of any gypsy sound.'

spelled <гад> (cf. ibid.: 199). The greatest difficulty in the opinion of the authors is the representation of unstressed vowels. In such a case they allow what they call etymological spelling. As an example they present three possibilities to spell a loan word from Tatar meaning 'work': <хезмети>, <хызмэти> or <хэзмэти> (all with the stress on the second syllable). To unify the diction of the first, unstressed vowel the authors decide for the spelling with <ы>, referring to the etymology of the word (cf. ibid.: 200). First ideas on a written form of Crimean Romani can be found already in Toropov (1999). He introduces them by means of a letter, written in Romani by a Crimean Rom, which is analyzed and corrected by him according to his own ideas about a Crimean Romani orthography (cf. Toropov 1999: 16f.).

3.2.2. Czech Republic, Slovakia
The second oldest attempt to script Romani in a Slavic-speaking country after Sergievskij & Barannikov (1938) dates back to the year 1969 and is based on the orthographies of Czech and Slovak. The publication of literary texts in Romani in the region started at the same time, around the end of the 1960s and the beginning of the 1970s.

In 1969, the Czechoslovak Romani organization *Svaz Cikánù-Romù* (SCR) was founded and started publishing the bulletin *Romano l'il*, which also contained texts in Romani from the third issue onwards. The chief editor, Andrej Pešta, developed an orthography for Slovak Romani (because of the numerical majority of the Slovak Roma and because all the productive Romani writers at that time used this dialect in their publications), the so-called SCR-orthography. Their basic principles were published in the journal *Romano l'il nevo* (cf. Hübschmannová 1993) and are used, with slight modifications, in numerous Romani publications (cf. Hübschmannová 1995: 193). The alphabet consists of the following letters: <Aa, Bb, Cc, Čč, Chch, Dd, Ďd', Dzdz, Džž, Ee, Ff, Gg, Hh, Chch, Ii, Jj, Kk, Khkh, Ll, L'l', Mm, Nn, Ňň, Oo, Pp, Phph, Rr, Ss, Šš, Tt, Ťt', Thth, Uu, Vv, Zz, Žž>. A peculiarity is the use of the acute: it does not, as in Czech and Slovak, represent long vowels, which originally do not exist in Romani, but the short form of a future or imperfect tense such as <kerás> 'I did' – so, morphology is represented in writing here. The palatalization of consonants is marked by a háček or an apostrophe.

Final devoicing is not represented in writing, e.g. <dad> [daːt] 'father'. The grapheme <x> should be avoided and <ch> should be written instead (cf. ibid.: 197). Regarding the acceptance of the project, Hübschmannová & Neustupný (1996: 100f.) come to a positive conclusion:

> Out of 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. [...] Overall we can say that [...] 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.

The diction was taken over with only minimal changes for the *Romsko-český a česko-romský kapesní slovník* (Hübschmannová, Šebková & Žigová 1991) which was published in the year of the split of Czechoslovakia. In the years after the political turnover a uniform version of Romani was fostered through journals and books and the number of publications in Romani was rising. However, the diction rather developed in a manner that Hübschmannová (1995: 196) calls "trial and error" than through language planning. Some authors applied the SCR-orthography, others spelled Romani spontaneously by ear. In the face of these developments, Hübschmannová & Neustupný (1996) plead for a "postmodern" and "polycentric" approach to writing Romani, based on already existing written texts and accepting variation. They are convinced that variation does not necessarily lead to difficulties in understanding but is rather beneficial to the acceptance of the endeavor by the speech community. The claim for unification takes a back seat. Matras (1999b: 99) views this "trial and error" principle as a simple result of pragmatical necessity, rather than a symbolic, elitist project.

Possibly, a few years later the need for a more uniform diction of Romani in former Czechoslovakia rose again, because in 2006 the first and so far only system of rules for writing Romani in independent Slovakia was published, written by a large collective of authors around Milena

Hübschmannová,[14] titled *Pravidlá rómskeho pravopisu.* It is the wish of the authors that their work may contribute to the preservation of the cultural and linguistic heritage of the Roma (cf. Hübschmannová et al. 2006: 9). It is supposed to be a useful assistance both for philologists and non-philologists and was developed on the basis of the variety of Romani spoken by 80% of the Romani population in Slovakia (cf. ibid.: 8) – i.e. the Northern Central dialect. The alphabet is, of course, Latin and consists of the following letters: <Aa, Bb, Cc, Čč, Čhčh, Dd, Ďd', Dzdz, Dždž, Ee, Ff, Gg, Hh, Chch, Ii, Jj, Kk, Khkh, Ll, L'l', Mm, Nn, Ňň, Oo, Pp, Phph, Rr, Ss, Šš, Tt, Ťt', Thth, Uu, Vv, Zz, Žž>. Loan words and proper names can also contain <Qq, Ww, Xx> or <Yy> (cf. ibid.: 18). The authors rely on the phonemic principle, which is, however, in special cases complemented by the morphological, etymological and the so-called interdialectal principle (for explanations see ibid.: 19f.). As Romani in contrast to Czech and Slovak does not know long vowels, the orthography does not contain long accents apart from a few exceptions (cf. ibid.: 21). In this project there is also only one <r> which is pronounced like the Slovak /r/, and aspirated consonants are written with an <h>: <čhon> [t͡ʃʰɔn] 'moon', <kham> [kʰam] 'sun', <phuv> [pʰuːf] 'earth', <thud> [tʰuːt] 'milk'. Final devoicing is not represented in writing and palatalization is marked with an apostrophe.

3.2.3. Macedonia

Romani has been used in Macedonia in a written form at least since the 1960s, and the first to conduct the experiment of creating an alphabet for Romani in Macedonia were Jusuf & Kepeski.[15] In 1973, they wrote a bilingual (Romani and Macedonian) grammar based on the two dialects Arli and Džambazi, which was, however, only published in 1980. At the

---

[14] Milena Hübschmannová (1933–2005), an Indologist from Prague, is presumed to be the founder of Romani studies in Czechoslovakia. Although she is not a Romni herself, she spoke Romani fluently and was awarded with many prizes for her commitment in matters of Roma.

[15] Šaip Jusuf (ca. 1933–2010) was a Rom from Skopje, a sports teacher with a diploma from the University of Belgrade (https://de.wikipedia.org/wiki/%C5%A0aip_Jusuf). Krume Kepeski (1909–1988), a *gajo*, was a linguist in Skopje (https://de.wikipedia.org/wiki/Krume_Kepeski).

same time, they introduced an orthography for this variety based on the Latin (!) alphabet, which they justify as follows:

> Авторите на оваа книга беа во дилема кои знаци да гиземат за обележување на гласовите што се слушаат кај Ромите во нашата земја; најпосле се решија за латиницата, бидејќи има Роми и во европски земји каде што народите со кои живеат се служат со латиница[16] (Kepeski & Jusuf 1980: 19).

Arli is, until today, the dialect that is spoken by the majority of Roma in Macedonia as well as in Serbia and Kosovo (this is the region Kepeski & Jusuf refer to) and that the Romani literature in the area is based on. For some sounds from this dialect which have no counterparts in Latin, special signs had to be found or Cyrillic letters had to be used: for example, the schwa is represented as <ä>. Friedman (1985) comments, that "the choice of yet another separate letter for Romani is not without justification", yet points to the fact that the schwa sound is so rare in the respective dialects that it was not really necessary to introduce a special grapheme for it. He concludes: "The problem of schwa in Romani dialects and the literary standard is thus clearly in need of further elucidation" (ibid.). The Romani alphabet consists of 32 graphemes (see fig. 2). Macedonian influence is visible in the letters <kj> (cf. mak. <ќ>) and <gj> (cf. mak. <ѓ>), which is also commented on by Friedman (1985: 58):

> According to RG [*Romani Gramatika*, annotation by AMM], the Romani sounds are closer to the Macedonian sounds represented by the Cyrillic letters ќ, ѓ than they are to the Serbo-Croatian sounds represented by the Cyrillic ћ, ђ and the Latin ć, đ. It is certainly the case that in the pronunciation of palatal or palatalized stops or affricates the various dialects of Romani often agree most closely with the pronunciation of the non-Romani languages or dialects with which they are in closest contact [...]. The dialects of RG are typical in this respect.

Sibilants are written with a háček and final devoicing is not represented in writing. As often, a difference between <x> and <h> is made, but: "Jusuf and Kepeski (1980) fail to make the distinction in practice, using

---

[16] 'The authors of this book found themselves in the dilemma which signs they should use to mark the sounds heard among the Roma in our country. In the end they decided for the Latin script, as there are Roma in European countries in which the people they live with use the Latin script.'

both <x> and <h> in the same roots, e.g. *xiv, hiv* 'hole', *xhor* 'depth' but *horadaripe* 'deepening' [...]" (Friedman 1995: 182). The aspirated consonants are, surprisingly, not explicitly named in the overview (cf. fig. 2), but elsewhere in the text one can find an explanation that they are spelled <h> in Latin and <x'> in Cyrillic, e.g. <than>/<тх'ан> 'place', <kham>/ <кх'ам> 'sun', <čhaj>/<чх'aj> 'girl', <phen>/<пх'ен> 'sister' (cf. Kepeski & Jusuf 1980: 23). /h/ is also spelled <x'> in the Cyrillic version, e.g. <hava>/<х'ава> 'to eat' (cf. ibid.). There is one grapheme for the velar (Latin and Cyrillic <x>) and one for the glottal fricative (Latin <h>, Cyrillic <x'>).

Friedman (1985: 59) also mentions that in Kepeski & Jusuf's system of rules, both <lj> and <l> can stand in front of front vowels, so that one can find both the spellings <ljil> and <lil> for 'book', to name just an example. According to him, a clear decision should be made here if /lj/ and /l/ are to be seen as separate phonemes. Regarding voicing, Friedman (1985: 59) attests the authors a relatively consistent approach:

> The orthography in RG is generally consistent in portraying underlying voiced and voiceless consonants in environments of neutralization, as indicated in the examples just given, although no explicit rules are stated, and occasional slips do occur, e.g. the spelling of *dat* for *dad*.

As in most Macedonian varieties of Romani, there is only one /r/, so that no differentiation in writing is necessary here. The alphabet is used, among others, in the *Makedonsko-romski i Romsko-makedonski rečnik* by Petrovski & Veličkovski (1998).

In 1992, a year after Macedonia's independence from Yugoslavia, a conference was held in Skopje with the aim of standardizing Romani and introducing the language as a subject in Macedonian schools (cf. Friedman 1995: 179).[17] Kurth comes to the conclusion that the standardization of Romani has made constant progress since the *Romani Gramatika*:

---

[17] The conference and its final document are elaborately presented in Friedman (1995). Furthermore, according to Kyuchukov (2009: 60), a second conference was held in Skopje in 2006 with Romani participants from Macedonia, Serbia, Bosnia and Bulgaria who discussed a standardization of Romani on a regional level.

PAŠAKJERDO DIKKJERIBA KO ROMANO-LATINSKO THAJ KIRILSKO HRAMONDO

romani	latinica	
A, a	A, a	А, а
Ä, ä	—, —	—, —
B, b	B, b	Б, б
C, c	C, c	Ц, ц
Č, č	Č, č	Ч, ч
Kj, kj (Ć, ć)	Ć, ć	Ќ, ќ
D, d	D, d	Д, д
Gj, gj (Đ, đ)	Đ, đ	Ѓ, ѓ
Dž, dž	Dž, dž	Џ, џ
E, e	E, e	Е, е
F, f	F, f	Ф, ф
G, g	G, g	Г, г
H, h	H, h	—, —
X, x	—, —	Х' х'
I, i	I, i	И, и
J, j	J, j	Ј, ј
K, k	K, k	К, к
L, l	L, l	Л, л
Lj, lj	Ll, lj	Љ, љ
M, m	M, m	М, м
N, n	N, n	Н, н
Nj, nj	Nj, nj	Њ, њ
O, o	O, o	О, о
P, p	P, p	П, п
R, r	R, r	Р, р
S, s	S, s	С, с
Š, š	Š, š	Ш, ш
T, t	T, t	Т, т
U, u	U, u	У, у
V, v	V, v	В, в
Z, z	Z, z	З, з
Ž, ž	Ž, ž	Ж, ж

Fig. 2: "Overview of the Romani, Latin and Cyrillic script" (Kepeski & Jusuf 1980: 20f.)

НАПОРЕДЕН ПРЕГЛЕД НА РОМСКОТО, ЛАТИНСКОТО
И КИРИЛСКОТО ПИСМО

ромско	латиница	кирилица
A, a	A, a	А, а
Ä, ä	—, —	—, —
B, b	B, b	Б, б
C, c	C, c	Ц, ц
Č, č	Č, č	Ч, ч
Kj, kj	Ć, ć	Ќ, ќ
D, d	D, d	Д, д
Gj, gj	Đ, đ	Ѓ, ѓ
Dž, dž	Dž, dž	Џ, џ
E, e	E, e	Е, е
F, f	F, f	Ф, ф
G, g	G, g	Г, г
H, h	H, h	Х' х'
X, x	—, —	Х, х
I, i	I, i	И, и
J, j	J, j	Ј, ј
K, k	K, k	К, к
L, l	L, l	Л, л
Ll, lj	Lj, lj	Љ, љ
M, m	M, m	М, м
N, n	N, n	Н, н
Nj, nj	Nj, nj	Њ, њ
O, o	O, o	О, о
P, p	P, p	П, п
R, r	R, r	Р, р
S, s	S, s	С, с
Š, š	Š, š	Ш, ш
T, t	T, t	Т, т
U, u	U, u	У, у
V, v	V, v	В, в
Z, z	Z, z	З, з
Ž, ž	Ž, ž	Ж, ж

Fig. 2 cont'd: "Overview of the Romani, Latin and Cyrillic script" (Kepeski & Jusuf 1980: 20f.)

151

> Die Konferenz von 1992 trug trotz Versuchen von politischer Instrumentalisierung durch alle Fraktionen dazu bei, die absolute Notwendigkeit einer konsistenten Orthographie aufzuzeigen. Die Einführung des Romischen als Unterrichtssprache auf der Grundstufe wird einen grossen [sic] Einfluss auf den Kodifikationsprozess ausüben, und zwar über die Grenzen Makedoniens hinaus (Kurth 2008: 55).[18]

One of the results was a new suggestion for an orthography, which took away parts of the scope Kepeski & Jusuf had left for the writers and strove for more uniformity and stricter rules. The new suggestion allowed that Romani in Macedonia could be written both with Latin and with Cyrillic letters: <Aa, Bb, Cc, Čč, Čh/čh, Dd, Dž/dž, Ee, Ff, Gg, Hh, Ii, Jj, Kk, Kh/kh, Ll, Mm, Nn, Oo, Pp, Ph/ph, Rr, Ss, Šš, Tt, Th/th, Uu, Vv, Žž> and <Аа, Бб, Цц, Чч, Чх/чх, Дд, Џџ, Ее, Фф, Гг, Хх, Ии, Jj, Кк, Кх/кх, Лл, Мм, Нн, Оо, Пп, Пх/пх, Рр, Сс, Шш, Тт, Тх/тх, Уу, Вв, Зз, Жж>.[19] Also in Macedonia, activists involved in Romani standardization speak out against an internationally normed orthography and prefer to follow the trend of regional or national orthographies. Although some publications in Macedonia have been published in Cortiade's orthography, the conference participants agreed on the practice of taking spontaneously produced Romani texts as an example, as it had been done in Czechoslovakia. This concerns, among others, the use of <dž> instead of <ʒ> for /d͡ʒ/. Cortiade's acute for marking sibilants was also rejected in favour of a háček (cf. Friedman 1995: 181f.). This orthography, in contrast to Kepeski & Jusuf (1980), does not contain a grapheme for the representation of the schwa: "In the rare instances of schwa in the Arlija dialect, the corresponding form in Džambaz or some other Romani dialect with a different vowel will be taken as the literary norm", says the document (cf. Friedman 1995: 183). In spite of all efforts to script and standardize Romani in Macedonia, Friedman (2005: 171) concludes:

> Aside from the orthography conference of 1992, norm selection is processing in Macedonia de facto rather than de jure. In this sense, the pro-

---

[18] 'The conference of 1992, in spite of attempts of political instrumentalization from all fractions, contributed to the demonstration of the absolute necessity of a consistent orthography. The introduction of Romani as a language of instruction on a basic level will have a great influence on the codification process, beyond the borders of Macedonia.'

[19] The order of the Cyrillic letters follows the Latin alphabet here.

cess of Romani standardization in Macedonia is following patterns seen for Romani in other countries which is to say that consensus is emerging through usage.

3.2.4. Bulgaria

Romani literature in Bulgaria began to develop in the 1950s (cf. Toninato 2014: 82), the first bilingual Bulgarian-Romani reader for children entitled *Romano ABC lil* was published in 1993 (cf. Kyuchukov 2009: 56). The Romani books that have been published for the use at Bulgarian universities play an important role in the process of standardizing an orthography for Romani in the country (cf. ibid.: 63f.).

As probably the majority of publications on Romani in Bulgaria were penned by Christo Kjučukov,[20] it is not surprising that both suggestions for scripting Romani in Bulgaria have been excogitated by him, both in cooperation with another researcher. The first orthography from the 1990s (Kjučukov & Yanakiev 1996; cf. also the summary in Kyuchukov 2009: 58) is notable in so far as it is also based on Latin script and additionally on English orthography. To script the schwa, the grapheme <w> was chosen, and palatalized vowels are written with a <y>: <ya, ye, yo, yu>. Sibilants are spelled as digraphs with an <h>: <chh, kh, ph, th>. This orthography, although designed for the Roma in Bulgaria, makes a very international impression as if it was influenced by English orthography and thus clearly delimits itself from the Bulgarian Cyrillic alphabet.

In the year 2000, a new suggestion for scripting Romani in Bulgaria was made (Hancock & Kjučukov 2000) which takes over many elements from the first version, but also makes some changes: the schwa is now represented as <y> and the palatalized vowels are spelled with <j> instead of <y>: <ja, je, jo, ju>. The English influence is reduced by the fact that sibilants are now written with a háček (<š, č, ž, dž>) as in earlier works by Hancock with an international orientation (cf. also the summa-

---

[20] Christo S. Kjučukov, born in 1962 in Provadija as Chjusein Selimov Kjučukov, is a specialist of Romani linguistics and education. He holds a Ph.D in General Linguistics from the University of Amsterdam and other academic titles (https://bg.wikipedia.org/wiki/Христо_Кючуков).

ry in Kyuchukov 2009: 58f.). It is hard to tell in how far the two projects have been applied in practice beyond Kjučukov's own publications.

3.2.5. Poland

The first known written Romani texts produced by Polish Roma date back to the woman poet Bronisława Wajs, known as Papusza (1910?–1987). Her poems have been edited since the beginning of the 1950s both in Romani and in Polish by Jerzy Ficowski; Papusza's original diction is spontaneous and intuitive. Until not long ago, only few Roma in Poland saw the necessity of using their language in a written form. Only within a short time in 2007/08 many publications in Romani were released in which the problem of orthography was solved in very different ways.

In the course of this development, Adam Bartosz[21] published the first systematical proposal for a Romani orthography in Poland (Bartosz 2009). The project is called *pisownia sulejowska* 'Sulejów diction' after the place where the people involved in the project gathered. Bartosz, who was in charge, describes it as the Polish reply to Hübschmannová's proposals for the Czech Republic and Slovakia (cf. Bartosz 2009: 158f.). He (ibid.: 164f.) emphasizes that the *pisownia sulejowska* is to be seen as a draft and basis for discussion, not as a final system of rules. It is based on the main dialects of Romani that are spoken in Poland and the influence of Polish orthography is clearly visible. Bartosz justifies this as follows:

> Pojawiła się potrzeba ujednolicenia zapisu, ale próby nawiązania do wcześniej ustalonych zasad napotkały niemały opór zainteresowanych. [...] Nie ma bowiem szansy na wprowadzenie w najbliższym czasie instytucjonalnych form nauczania pisowni romskiej (brak nauczycieli, systemu nauczania, zapotrzebowania samego środowiska na takie nauczanie etc.). W takiej sytuacji należy się zdecydować na zastosowanie pisowni w oparciu o alfabet polski (Bartosz 2009: 160f.).[22]

---

[21] Adam Bartosz is a *gajo* but speaks Romani fluently. He is a productive ethnographer and was director of the district museum in Tarnów until 2012. The museum houses the largest exhibition on Romani culture in Poland.

[22] 'A necessity of unifying the orthography had appeared, but attempts to rely on earlier defined rules evoked resistance among the people involved. [...] Namely, there is no chance to introduce institutionalized forms of teaching a writing system for Romani in

The diction refers to the tendencies that can be seen in spontaneous Romani text production in Poland. It consists of the following graphemes: <Aa, Bb, Cc, Ćhćh, Dd, Dźdź, Ee, Ff, Gg, Hh, Chch, (Xx), Ii, Jj, Kk, Khkh, Ll, Łł, Mm, Nn, Ńń, Oo, Pp, Phph, Rr, Ss, Śś, Tt, Thth, Uu, Ww, Vv, Yy, Zz, Źź>. Aspirated consonants („[j]est to bowiem istotna cecha języka *romani*, którą w pisowni należy zachować"[23] – ibid.: 164) are spelled with an <h>, e.g. <ćhaj> 'girl', <kher> 'house', <phabaj> 'apple', <them> 'place, country'. The uvular fricative can be spelled <ch> or <x>, however, Bartosz emphasizes the difference between /x/ and /h/, which are separate phonemes (ibid.: 163). A special feature of the *pisownia sulejowska* is the grapheme <ł> for the labialized velar approximant [w], which unmistakeably demonstrates Polish influence, for example in <łoło> [ˈwɔwɔ] 'red' instead of <lolo> [ˈlɔlɔ] in other Romani dialects (cf. ibid.: 162). The diction also reflects a more palatalized pronunciation of certain consonants – also under the influence of Polish – which is visible in the graphemes <ć, ćh, dź, ń, ś, ź> (cf. ibid.: 163). However, other rules prevail for them in the *pisownia sulejowska* than in Polish orthography: In the latter, these graphemes could never be followed by an <i>, in the former this is absolutely possible, see Bartosz's examples <ćaćipen> 'truth' or <podźi> 'skirt' (ibid.: 161). This makes sense because Romani – in contrast to Polish – has words containing the phoneme combinations /si/, /ci/ etc. They are spelled <si>, <ci> in the *pisownia sulejowska*, thus <sikaweł> is pronounced [sikavˈɛw] instead of [ṣikavˈɛw]. There is only one realisation of /r/ in the Polish Romani dialects, so only one grapheme is necessary. To spell Polish loan words, it is also allowed to introduce other graphemes from the Polish alphabet and Bartosz pleads for their unchanged adoption (cf. ibid.: 164). The proposal was received positively by the speech community and Mirga (2009) explicitly uses it for his *Słownik romsko-polski*.

---

foreseeable time (lack of teachers, an education system, the demand even for an environment for such a form of education etc.). In such a situation it is necessary to decide for the use of a script on the basis of the Polish alphabet.'

[23] 'Since this is an essential feature of Romani which must be preserved in writing'.

4. Conclusion

The aim of the present article was to gather information about all projects of creating an alphabet and an orthography for Romani in Slavic-speaking countries and to analyze and compare them against the theoretical background of Smalley (1964) and Coulmas (1989). It is important to notice that these are the attempts from this area known to me – it cannot be excluded that there are even more that have not found wider dissemination. But even if not exclusively all projects could be found, the ones presented here demonstrate the main tendencies very well.

The ten examples from seven countries that have been presented here demonstrate that the trend for regional or national instead of international approaches to finding a script and an orthography for Romani (cf. Matras 2005) has been lasting for several decades. They differ as to whether the authors propose a solution for all the important Romani dialects in their country taken together (e.g. Poland), only for the most important one(s) (e.g. Macedonia) or even make several different suggestions for different Romani dialects in the country (e.g. Russia).

It has become clear that the proposals in most cases are based on the script and alphabet of the respective country's majority language, so that one can speak more of an approximation than of a demarcation from it. This approximation should, however, be interpreted more pragmatically than emotionally, because some authors clearly emphasize that their alphabet is something of its own even though the same script is used as in the country's majority language. This is underlined by the special graphemes and diacritic signs that can be found in every alphabet. Interesting exceptions in this respect are Macedonia and Bulgaria. Although one would expect a fallback on Cyrillic in these cases, the authors decide for Latin with regard to the wider international range of their alphabets.

The dominant principle presented by the authors – both Roma and *gaje* are among them – is the phonemic one. Sometimes the authors make concessions to other principles or are not totally consistent in its application, but generally this is the prevailing and aspired way.

In how far the scripting suggestions have been accepted by the respective speech community is probably the question that is most difficult to answer. What can be proven is that most of them have later been used in

publications such as dictionaries, grammars or textbooks, i.e. in an academic context. In order to tell how many Roma use them in their private communication, in internet blogs or in non-academic publications, an analysis of a wider range would be necessary.

Presently, the process of trying to find adequate scripts and alphabets for Romani in different Slavic-speaking countries is still going on. Very currently, the first Romani-Montenegrin dictionary was published in Podgorica (Demir, Durmiš & Demir 2015), and in Slovenia a project for a regional standardization was launched already in 2003 (Antauer, Živa & Peršak 2003), the outcome of which is however unknown to me. In Croatia, for example, there is no elaborated proposal for scripting Romani as in the countries analyzed above, but basic considerations in this direction can be found for example in the *Romsko-hrvatski i hrvatsko-romski rječnik* by Kajtazi (2008: 17). The same applies for Đurić's *Gramatika romskog jezika* (cf. Đurić 2005) as well as for other countries that have not been explicitly named in the present paper. The efforts have not come to an end yet.

Bibliography

ANTAUER, ŽIVA & PERŠAK (2003) – Živa Antauer, Janez Krek & Mateja Peršak, *Projekt standardizacije jezika Romov v Sloveniji in vključevanje romske kulture v vzgojo in izobraževanje.* In: Vera Klopčič & Miroslav Polzer (eds.), *Evropa, Slovenija in Romi. Zbornik referatov na mednarodni konferenci v Ljubljani, 15. Februarja 2002.* Ljubljana: Inšt. za Narodnostna Vprašanja, 131–141.

BARTOSZ (2004) – Adam Bartosz, *Nie bój się Cygana. Na dara Romestar.* Sejny: Pogranicze.

BARTOSZ (2009) – Adam Bartosz, *Propozycja zapisu języka romani – Pisownia sulejowska.* «Studia Romologica» 2: 153–166.

BUßMANN (2008) – Hadumod Bußmann, *Lexikon der Sprachwissenschaft.* 4th ed. Stuttgart: Kröner.

COULMAS (1989) – Florian Coulmas, *The Writing Systems of the World.* Oxford—New York: Blackwell.

COURTHIADE (2012) – Marcel Courthiade, *„Alfabet odpowiedni dla języka cygańskiego oparty na naukowych podstawach" – o pisowni języka rromani.* Tarnów: Muzeum okręgowe.

CVETKOV (2001) – Георгий Николаевич Цветков, *Романэ ворби. Цыганско-русский и русско-цыганский словарь (ловарский диалект). Около 4000 слов.* Москва: ООО «НИЦ Апостроф».

DEMETER & DEMETER (1990) – Роман Степанович Деметер & Петр Степанович Деметер, *Цыганско-русский и русско-цыганский словарь (кэлдерарский диалект). 5300 слов. Под ред. Л.Н. Черенкова*. Москва: Русский язык.

DEMIR, DURMIŠ & DEMIR (2015) – Ljatif Demir, Nevsija Durmiš & Fatime Demir, *Crnogorsko-romski i Romsko-crnogorski rječnik*. Podgorica: Zavod za udžbenike i nastavna sredstva.

DJURIĆ (2002) – Rajko Djurić, *Die Literatur der Roma und Sinti*. Berlin: Ed. Parabolis.

ĐURIĆ (2005) – Rajko Đurić, *Gramatika e rromane čhibaki. Граматика ромског језика*. Beograd: Откровење.

FRIEDMAN (1985) – Victor A. Friedman, *Problems in the Codification of a Standard Romani Language*. In: Joanne Grumet (ed.), *Papers from the Fourth and Fifth Annual Meetings* (Gypsy Lore Society Publications 2). New York: Gypsy Lore Society, North American Chapter, 56–75.

FRIEDMAN (1995) – Victor A. Friedman, *Romani standardization and status in the Republic of Macedonia*. In: Yaron Matras (ed.), *Romani in contact. The history, structure and sociology of a language* (Amsterdam Studies in the Theory and History of Linguistic Science 126, Series IV – Current Issues in Linguistic Theory). Amsterdam—Philadelphia: John Benjamins, 177–188.

FRIEDMAN (2005) – Victor A. Friedman, *The Romani Language in Macedonia in the Third Millennium: Progress and Problems*. In: Barbara Schrammel, Dieter W. Halwachs & Gerd Ambrosch (eds.), *General and Applied Romani Linguistics. Proceedings from the 6th International Conference on Romani Linguistics*. München: Lincom Europa, 163–173.

HANCOCK & KJUČUKOV (2000) – Ian Hancock & Xristo Kjučukov, *Angluni vorba. Atanasov-Ramar, Toško, Čhordo kamipe*. Austin: Romanistan Publ.

HEINSCHINK & CECH (2013) – Mozes F. Heinschink & Petra Cech, *Die Sprache der Roma – Grundzüge der Romani Čhib und Perspektiven für den Unterricht*. In: Sabine Hornberg & Christian Brüggemann (eds.), *Die Bildungssituation von Roma in Europa* (Studien zur International und Interkulturell Vergleichenden Erziehungswissenschaft 16). Münster etc.: Waxmann, 53–87.

HÜBSCHMANNOVÁ (1993) – Milena Hübschmannová, *Hlavné zásady romského pravopisu*. «Romano l'il nevo» 3: 88–99.

HÜBSCHMANNOVÁ (1995) – Milena Hübschmannová, *Trial and error in written Romani on the pages of Romani periodicals*. In: Yaron Matras (ed.), *Romani in contact. The history, structure and sociology of a language* (Amsterdam Studies in the Theory and History of Linguistic Science 126, Series IV – Current Issues in Linguistic Theory). Amsterdam—Philadelphia: John Benjamins, 189–205.

HÜBSCHMANNOVÁ, ŠEBKOVÁ & ŽIGOVÁ (1991) – Milena Hübschmannová, Hana Šebková & Anna Žigová, *Romsko-český a česko-romský kapesní slovník*. Praha: Statní pedagogické nakladatelství.

HÜBSCHMANNOVÁ & NEUSTUPNÝ (1996) – Milena Hübschmannová & Jiří Neustupný, *The Slovak-and-Czech dialect of Romani and its standardization*. «International Journal of the Sociology of Language» 120: 85–109.

HÜBSCHMANNOVÁ ET AL. (2006) – Milena Hübschmannová et al., *Pravidlá rómskeho pravopisu. S pravopisným a gramatickým slovníkom*. Bratislava: Štátny Pedagogický Ústav.

IGLA (1991) – Birgit Igla, *Probleme der Standardisierung des Romani*. In: James R. Dow & Thomas Stolz (eds.), *Akten des 7. Essener Kolloquiums über „Minoritäten / Sprachminoritäten" vom 14.–17.6.1990 an der Universität Essen* (Bochum–Essener Beiträge zur Sprachwandelforschung 10). Bochum: Brockmeyer, 75–90.

KAJTAZI (2008) – Veljko Kajtazi, *Romano-kroacijako thaj kroacijako-romano alavari. Romsko-hrvatski i hrvatsko-romski rječnik*. Zagreb: «Kali Sara» i Odjel za orijentalistiku Hrvatskoga filološkog društva.

KENRICK (1981) – Donald Kenrick, *The development of a standard alphabet for Romani*. «The Bible translator» 32, 2: 215–219.

KEPESKI & JUSUF (1980) – Круме Кепески & *Šaip Jusuf, Romani Gramatika. Ромска граматика*. Скопје: ООЗТ Книгоиздателство Наша Книга.

KJUČUKOV & YANAKIEV (1996) – Христо Кючуков & Мирослав Янакев, *Ромският път. О Романо Дром*. София: Тилиа.

KOCHANOWSKI (1995) – Vania de Gila Kochanowski, *Romani Language Standardization*. «Journal of the Gypsy Lore Society» 5, 2: 97–108.

KURTH (2008) – Gérald Kurth, *Identitäten zwischen Ethnos und Kosmos. Studien zur Literatur der Roma in Makedonien*. Wiesbaden: Harrassowitz.

KYUCHUKOV (2009) – Hristo Kyuchukov, *Standardization of Romani in Bulgaria*. In: Id. (ed.), *"A Language without Borders..." The International Romani Language Conference in Stockholm and Uppsala, 7–9 January 2007* (Endangered languages and cultures 5). 2nd ed. Uppsala: Centre for Multiethnic Research, 56–65.

MACHOTIN (1993) – Джура Махотин, *Аджутипе прэ романи чиб. Пособие по языку романи*. Тверь: без изд.

MATRAS (1997) – Yaron Matras, *Probleme und Chancen der Verschriftung der Sprache der Roma und Sinti*. In: Heiko Balhorn & Heide Niemann (eds.), *Sprachen werden Schrift. Mündlichkeit – Schriftlichkeit – Mehrsprachigkeit* (Libelle: Wissenschaft lesen und schreiben 7). Lengwil am Bodensee: Libelle, 110–116.

MATRAS (1999a) – Yaron Matras, *Writing Romani: The Pragmatics of Codification in a Stateless Language*. «Applied Linguistics» 20, 4: 481–502.

MATRAS (1999b) – Yaron Matras, *Sprachplanung und Spracheinstellung im Romanes*. In: Dieter W. Halwachs & Florian Menz (eds.), *Die Sprache der Roma. Perspektiven der Romani-Forschung in Österreich im interdisziplinären und internationalen Kontext*. Klagenfurt: Drava, 95–111.

MATRAS (2002) – Yaron Matras, *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.

MATRAS (2005) – Yaron Matras, *The future of Romani: Toward a Policy of Linguistic Pluralism*. In: *errc – european roma rights centre*. www.errc.org/cikk.php?cikk=2165 (last access 20.01.2017).

MATRAS (2014) – Yaron Matras, *I Met Lucky People. The Story of the Romani Gypsies*. London: Allen Lane.

MIRGA (2009) – Jan Mirga, *Słownik romsko-polski*. Tarnów: Romskie Stowarzyszenie Oświatowe Harangos, Muzeum okręgowe.

ONG (1982) – Walter J. Ong, *Orality and literacy. The Technologizing of the Word*. London—New York: Methuen.

Petrovski & Veličkovski (1998) – Trajko Petrovski & Bone Veličkovski, *Makedonsko-romski i Romsko-makedonski rečnik. Makedonsko-romano i Romano-makedonsko alavari.* Skopje: Vorldbuk.

Piasere (1985) – Leonardo Piasere, *Mare Roma. Catégories humaines et structure sociale. Une contribution à l'éthnologie tsigane.* Paris: Paul-Henri Stahl.

Sergievskij (1931) – Максим Владимирович Сергиевский, *Цыганский язык. Краткое руководство по грамматике и правописанию.* Москва: Центриздат.

Sergievskij & Barannikov (1938) – Максим Владимирович Сергиевский & Алексей Петрович Баранников, *Цыганско русский словарь.* Москва: Гос. изд-во иностранных и национальных словарей.

Smalley (1964) – William A. Smalley, *Dialect and Orthography in Gipende.* In: Id. et al. (eds.), *Orthography Studies: Articles on New Writing Systems.* London—Amsterdam: United Bible Societies, North-Holland Publ. Co., 138–144.

Šapoval (2007) – Виктор Васильевич Шаповал, *Самоучитель цыганского языка (русска рома: севернорусский диалект). Учебное пособие.* Москва: АСТ Астрель Хранитель.

Toninato (2014) – Paola Toninato, *Romani Writing. Literacy, Literature and Identity Politics.* New York—London: Routledge.

Toropov (1999) – Вадим Германович Торопов, *История изучения цыганского языка в России.* В: Надежда Георгиевна Деметер, *Цыгане. Сборник статей* (Серия Новые исследования по этнологии и антропологии). Москва: Институт этнологии и антропологии им. Н. Н. Миклухо-Маклая РАН, 16–26.

Toropov (2003) – Вадим Германович Торопов, *Словарь языка крымских цыган.* Москва: РАН.

Toropov & Gumeroglyj (2013) – Вадим Германович Торопов & Павел Борисович Гумероглый, *Опыт составления алфавита для записи устной речи крымских цыган.* В: Кирилл Александрович Кожанов, Софья Алексеевна Оскольская & Александр Юрьевич Русаков (ред.), *Цыганский язык в России. Сборник материалов Рабочего совещания по цыганскому языку в России. Санкт-Петербург, 5 октября 2012г.* Санкт-Петербург: РАН, 199–228.

Wlislocki (1994) – Heinrich von Wlislocki, *Zur Ethnographie der Zigeuner in Südosteuropa. Tsiganologische Aufsätze und Briefe aus dem Zeitraum 1880–1905.* Ed. by Joachim S. Hohmann (Studien zur Tsiganologie und Folkloristik 12). Frankfurt am Main etc.: Lang.