PRZEMYSŁAW FAŁOWSKI
Jagiellonian University in Krakow
przemyslaw.falowski@uj.edu.pl

WORDS OF ROMANI ORIGIN IN THE CZECH AND CROATIAN LANGUAGES

Abstract

The paper is an attempt to compare words of Romani origin in the Czech and Croatian languages on the basis of two contemporary lexicographic sources. The main objective is to confirm the thesis regarding the presence of words with the same Romani etymon in both languages as well as to provide the semantic charactersitics of the analysed lexemes. The paper also presents information about the frequency of the words of Romani provenience in the Czech and Croatian languages that were collected using Internet corpora of both tongues.

1. Introduction

Words of Romani origin have been present in Czech and Croatian area for many years. They have long constituted a component of the lexis, namely slang in both languages, but they were primarily connected with the criminal world and had a character of cryptolects. It appears that Romani words, thanks to their obscurity, were ideal for concealing the content of the communications between, e.g.: vagabonds, beggars, card cheats or thieves. Interestingly, in the case of both languages these very old borrowings are still preserved in some sociolects (especially those that require secrecy), and moreover, some have even penetrated into the general colloquial lexis).[1]

---

[1] The term "general colloquial lexis" has no straightforward equivalent in either of the two tongues, hence it needs to be treated to a large extent as somewhat arbitrary. It results from the specific stratification of the Czech and Croatian languages. A special place in the stratification of the Czech language is occupied by the so-called *obecná čeština*, which is somewhat imprecisely and ambiguously defined as a language variety used in Bohemia and western Moravia as the so-called "third standard" (cf. Chloupek 2003: 51). Another popular term in Czech linguistics is that of

This phenomenon is puzzling, particularly in the context of the Polish language, in which words of Romani origin are in fact missing.[2] A further issue which needs to be carefully examined is the presence of words of Romani origin with the same etymon, which mostly also have the same meaning, in the lexical material of the Czech and Croatian languages.[3] It is of great interest, especially if we consider that we are dealing with languages which are direct neighbours.

Thus, the objective of this paper, apart from discussing the semantic characteristics of the Romani borrowings, will be to demonstrate their similarities within the lexis of both languages. A final task will be to appraise the frequency of the words of Romani origin by using Internet corpora of both tongues.

1.1. Information concerning the sources

The source of the analysed material were two contemporary dictionaries that cover what is broadly understood as non-standard lexis: the third edition of the Czech lexicon, entitled *Slovník nespisovné češtiny*, published in 2009 (Hugo 2009) and the Croatian *Rječnik hrvatskoga žargona* published in 2001 (Sabljak 2001).

The *Slovník nespisovné češtiny* is the outcome of a large-scale group research project concerning non-standard Czech lexis led by Jan Hugo, in which both Czech linguists and experts from various academic domains and other walks of life were involved. The dictionary not only makes use of all the Czech dictionaries and studies that focus on the non-literary Czech lexis, but also works that describe the character of this type of vocabulary in other languages. A considerable proportion of the material was obtained by traditional means using transcripts of recordings and questionnaires, although the most recent lexis also came from various Internet sources (cf. Hugo 2009: 23).

---

běžná mluva (or *běžně mluvený jazyk*), denoting a primarily spontaneous spoken variety of the national language commonly used in everyday communication. More information regarding the above issue can be found *i.a.* in the studies by Petr Sgall and Jiří Hronek (Sgall, Hronek 1992), Maria Krčmova (1997) and Jan Chloupek (2003). Refering to the variety of language used for everyday communication, Croat linguist Josip Silić uses the term *razgovorni jezik* (or *razgovorni stil*), which, as he believes, is not uniform, but which varies depending on the region in which it is spoken as well as the social group and the educational background of its users, and their origin (cf. Silić 1997: 488). At present, the spoken variety of Croatian is affected by strong regional variation resulting in the separation of four main regiolects: Kajkavian from the centre in Zagreb, Kvarner in Istria and Rijeka, Dalmatian (Split, Dubrovnik) as well as Slavonian (Osijek, Vukovar) (Oczkowa 2005: 90).

[2]   Cf., among others, the description of Polish sociolects in the article by Grabias, *Środowiskowe i zawodowe odmiany języka – socjolekty* (= *Social and professional varietes of the language – sociolects*) in which Hebrew, Yiddish, Greek, French, Ukrainian and Belarusian are mentioned as the main sources of the lexis of closed Polish social circles, with a complete absence of Romani examples (Grabias 2001: 244–245). A perfunctory analysis of the *Słownik tajemnych gwar przestępczych* (= *A dictionary of secret criminal slangs*) by Stępniak (1993), on the other hand, demonstrated the presence of three forms that were definitely of Romani provenience.

[3]   The Croatian language, regarded in the 20th century as a western variety of the Shtokavian dialect in the common Serbo-Croatian language, was not the only one in which Gypsy words were present. This is confirmed by the presence in the Serbian lexicon *Beogradski frajerski rečnik* by Petrit Imami (2007) of more than 100 lexemes of Romani provenience; cf. also the Gypsy words included in Amela Šehović's study of colloquial Bosnian (Šehović 2009).

The third edition of the lexicon *Slovník nespisovné češtiny* that is used in this paper contains ca. 17000 entries connected with former and present-day non-standard Czech language – these are elements of argot, lexemes derived from various types of slang (youth, criminal, military, police, etc.), occupational lects, vulgarisms, as well as words coming from sociolects and to some extent also from rural dialects. Nearly every entry is provided with an example of its use, sourced from classical as well as contemporary literature and the Internet. A considerable number of entries are provided with qualifiers and information concerning their origin, drawn primarily from various etymological dictionaries.

The *Rječnik hrvatskoga žargona* written by Tomislav Sabljak in 2001 is a considerably extended and updated version of a lexicon entitled *Šatra (rječnik šatrovačkog govora)* (published by Sabljak in 1981). In the *Rječnik hrvatskoga žargona* ca. 26 000 words and phraseologisms from all over Croatia were collected (by comparison, the *Rječnik šatrovačkog govora* contains only 5000 entries) and these are very varied. The author presents the situation regarding present-day Croatian slang and quotes examples from the so-called old jargon. There are numerous elements of former and contemporary colloquial lexis, regionalisms and elements of sociolects from Zagreb, Split and Osijek as well as lexemes coming from various types of slang. A small number of entries are provided with quotations, primarily from belles-lettres, classical and contemporary literature – in the case of certain an attempt to provide the etymology is indicated. Although a proportion of the entries are preceded by qualifiers describing the character of a given word, these are not as precise as in the Czech dictionary, as exemplified, for instance, a lack of data concerning some regionalisms.

It is worth mentioning the short *Introduction* to the 1981 edition, which clarifies the position of the *šatrovački govor* compared to the remaining varieties of the Croatian language. It also contains an overview of the ways non-standard lexis can be created as well as a list of publications concerning the *šatrovački govor* and related linguistic codes.

1.2. Information concerning the language material

The analysed language material consists of 404 Czech[4] and 195 Croatian lexemes which are loans from Romani, or possibly derivatives from those loans (among others, there are also compounds and products of contamination). These represent various parts of speech: in both languages the most frequent are nouns (Czech 260, Croatian 147), followed by verbs (Czech 86, Croatian 46). Moreover, in the Czech material one may also find a considerable number of adjectives (Czech 38, Croatian 2), as well as the occasional numeral (Czech 8) and adverb (Czech 7), in addition to 1 preposition and 4 words belonging to other parts of speech.

---

[4] Each separate word-formative derivative as well as each separate meaning of a given word is recognised as a separate lexeme here. As a result, the number of lexemes is much higher than the number of entries excerpted from the lexicographic sources. The number of Romani etymons which constitute the basis for borrowings in both languages is also important: the Czech lexemes were formed from 163 Romani etymons, while the Croatian – from 43.

The analysed lexemes represent non-standard varieties of the two languages that differ socially, professionally or territorially, which in the Czech material is specified by means of suitable qualifiers included in the dictionary. The Czech Gypsy words mostly belong (146 lexemes) to the so-called *mluva světských*, that is the language of travelling circus performers or comedians with no permanent place of residence, who are identified by certain researchers together with the ethnic Romani population (cf. Hugo 2009: 24–25). We should also mention 31 words which belong to both *mluva světských* and prison slang, which testifies to the penetration of many forms into groups of a similar social status. Words marked with the qualifier *vězeň.* (prison slang) appear in the Czech material 72 times, whereas the 19–20th c. argot is represented by 46 lexemes. We also include examples of criminal slang (26), youth slang (10) and police slang (1). 12 lexemes represent the sociolect of Brno, with 1 example taken from the Ostrava sociolect; there are also lexemes described as vulgar (7) and scornful (1).

The Croatian material has not been analysed in such detail as far as the typology of the lexis is concerned: we have 2 regionalisms, that is 2 words taken from the sociolect of Split and one Kajkavian example.

1.3. Research methods

For the etymological analysis of the Czech lexis, apart from the third edition of the *Slovník nespisovné češtiny* (Hugo 2009),[5] the Czech etymological dictionaries by Jiří Rejzek (2001) and Josef Holub and Stanislav Lyer (1968), as well as a paper by Karel Kamiš (1998) were also used. In the study of the Croatian lexis the etymological dictionary by Petar Skok (1971–1974), the dictionaries of foreign words by Bratoljub Klaić (2001) and Vladimir Anić and Ivo Goldstein (2000), the previously mentioned dictionary by Petrit Imami (2007), and most importantly a dissertation on Romani vocabulary by Rade Uhlik (1954) proved helpful.

In the description of the Romani borrowings, in the case of nouns the semantic aspect of the borrowed words was considered, based mainly on the division of the colloquial nouns into 23 semantic categories as proposed by Danuta Buttler. The Polish linguist enumerates the following categories: "the human being", "physical activities," "parts of the human body," "food," "clothing," "states and characteristics of the human being," "entertainment," "psychological processes of the human being," "situations and events taking place in the human community," "house tools and equipment," "shopping and money," "home," "study," "means of transportation," "social institutions," "animals," "work," "literature and press," "substances," "politics," "weather conditions," "measures," and "plants" (Buttler 1978: 37–45). This classification has been slightly modified for lexis under analysis and some categories have been changed. As a result 24 classes have been distinguished: "names of persons," "human activities", "parts of the body," "food," "clothing," "features and characteristics of the human being," "mental and physical

---

[5]   Most entries in the Czech dictionary had been thoroughly analysed in terms of genetics, the work being based on Czech etymological dictionaries and other academic studies. Thus, in the present paper most etymologies will be taken directly from the *Slovník nespisovné češtiny* and will be marked with the abbreviation Hugo 2009.

states of the human being," "music," "stimulants," "situations and events taking place in human communities," "tools and equipment," "furniture and furnishings," "money and shopping," "study," "means of transportation," "animals," "literature and press," "substances and materials," "weather conditions," "measures," "sport," "features of objects, situations and phenomena," "names of places and institutions," and "other."

A very important issue is the actual presence of words of Romani origin in contemporary spoken and written texts. The frequency of specific Romani borrowings in both these Slavic languages has been analysed using the Internet corpora: of the Czech language (*Český národní korpus*) and the Croatian language (*Hrvatski nacionalni korpus*).

2. The historical outline

Although the Roms appeared in Central Europe as early as 1417, and in Czech territory (in Prague) by 1419, at the beginning their contacts with the local population were to a large extent limited. These relations only developed in the 19th c. and concomitantly from the second half of the 19th c. the number of Romani words to enter the Czech argot increased (Hugo 2009: 19, 26). That words of Romani origin appeared in greater numbers is seen in the 1902 work by Karel Juda *Tajná řeč (hantýrka) zlodějů a šibalů*, and in the 1914 dictionary by František Bredler *Slovník české hantýrky (tajné řeči zlodějské)* in which they constitute a significant part of the lexis (Hugo 2009: 20). The importance of the influence of Romani words on the Czech argot at the beginning of the 20th is highlighted by František Oberpfalcer in his 1934 study, as he placed them in third position as a source of vocabulary (after German and Yiddish) for this language variety. The Czech linguist notes a considerable number of words of Romani provenience in the Czech argot, at the same time drawing attention to their distorted and semantically modified form in the borrowing language (Oberpfalcer 1934: 333). As regards the more recent works concerning the current state of Romani borrowings, it is worth citing the 1998 paper by Karel Kamiš, in which the author notes many argot words of Romani origin which had in the past been borrowed over by the Czech language. Additionally, he stresses a very important fact, namely that Romani words have never appeared in the literary variety of the Czech language (*spisovná čeština*), and in recent years they have also stopped penetrating into the non-literary varieties of Czech[6] (Kamiš 1998: 126–129).

In the Croatian linguistic tradition the Roms were believed to be the creators of a specific variety of language named *šatrovački govor* or *šatra*, which differed from all the other known varieties and used to date in Croatia. Tomislav Sabljak stresses the fact that it was the secret language of a group of people, a system of secret signs and ciphers inaccessible to other persons and groups[7] (Sabljak 2001: 5). It appears, however, that

---

[6] Kamiš's (1998: 129) statement that words of Romani origin "are hardly present in the language of non-Roms," however, is debatable.

[7] Later the meaning of the term *šatrovački jezik* was significantly extended, referring to a broader group of secret languages, whereas in recent years, under the influence of the American term *slang*, it also started to mean various kinds of sociolects (Fink 2003: 79).

instead of attributing the authorship of *šatrovački govor* to the Roms it would be better to state that *šatrovački govor* and the language of the Roms developed in close symbiosis, thus influencing each other. Such a position is taken by, among others, Rade Uhlik, the author of a comprehensive 1954 dissertation about the Romani influence on *šatrovački govor* in Bosnia and Hercegovina, who noted "*šatrovačkim govorom provejava tipičan ciganski duh ili, kao što neki smatraju, da je ciganski jezik nadahnut šatrovačkim duhom*" (Uhlik 1954: 6). Rade Uhlik also points to the connections between the Romani and the criminal communities as well as to the specific "Gypsy stylistics" which influenced the plasticity of *šatra* and other jargons (Uhlik 1954: 6). Interestingly, since the time of this pioneering research hardly any serious work on the Romani influence on the language of the Slavic majority have appeared in Serbia or Croatia.[8]

3. Semantic and word-formative analysis

3.1. Formal and semantic similarity of Romani borrowings

One of the most significant results of this comparative analysis is the detection of the presence in the analysed material of a considerable number of Czech and Croatian words with the same Romani etymon, the same meaning and a similar form. Approximately 20 such pairs of words may testify to the penetration of a similar category of Romani vocabulary into both languages. Specific examples are listed below:

- Czech *bul*, *bulovnice* sec.[9] 'bottom,' Croatian *bulj, bulja, buljina* 'ditto' < Rom. *bul* 'ditto' (Hugo 2009, Uhlik 1954: 10);
- Czech *čokl/čukl* derog. 'dog', Croatian *đukela* 'dog (large)' < Rom. *džukel* 'ditto' (Hugo 2009, Kamiš 1998: 128, Uhlik 1954: 15);
- Czech *čórnout* 1. 'to steal' 2. sec. 'to take', Croatian *ćornuti* 'to steal' < Rom. *čor* 'thief,' *čorel* 'to steal' (Hugo 2009; Kamiš 1998: 128; Uhlik 1954: 12), *te čorés* 'to steal' (Horbač 2006: 373);
- Czech *chalovat* sec. 'to eat', Croatian *halisati* 'ditto' < Rom. *chal* 'ditto' (Hugo 2009; Uhlik 1954: 16), Rom. *te cha(s)* 'to eat'; *chalo* 'eaten' (Horbač 2006: 384);
- Czech *káro, kár* prison sl. 'male genitals,' Croatian *kar* 'ditto' < rom. *kár/kar/ker* 'ditto' (Hugo 2009; Kamiš 1998: 129; Uhlik 1954: 19; Horbač 2006: 377);
- Czech *love/lóve* Brno sociolect, 'money', Croatian *lova* 'ditto' < Rom. *love/lóve* 'ditto' (Hugo 2009; Kamiš 1998: 129; Uhlik 1954: 21);
- Czech *mindža* vulg. 1. 'woman, stresses stupidity or only lack of affection' ← 2. 'female genitals,' Croatian *mindža* 1. 'woman, girl' ← 2. 'female genitals' < Rom. *mindž/minž* 'female genital' (Hugo 2009; Kamiš 1998: 129; Uhlik 1954: 24);
- Czech *páň/pháň* sec. prison sl. 1. 'water' 2. 'river', Croatian *panija* 'water' < Rom. *pani/páni/panji* 'water'[10] (Hugo 2009; Uhlik 1954: 25).

---

[8]   Nor are they mentioned by Željka Fink (2003) carrying out a detailed comparative analysis of old and present-day *šatrovački govor* in Croatia.

[9]   sec. = secular. See *mluva světských* in 1.2. Information concerning the language material.

[10] It is worth emphasising that these two lexemes also appear in the dictionary by Stępniak (1993) in Polonised forms: *mindzia mała* 'female genital,' *pani* 'water'.

In the analysed material it is also possible to find Czech-Croatian word pairs with the same Romani etymon, but which differ somewhat semantically or possess additional meanings, e.g.:

- Czech *kúlo/khúlo* sec. prison sl. 'excrements, dung,' *kulový* vulg. 'dung, shit, refusal,' Croatian *kulana* 'prison'[11] < Rom. *khul/khulo* 'excrement' (Hugo 2009; Uhlik 1954: 20);
- Czech *lil* 1. sec. 'permit, evidence, document' 2. prison sl. 'letter,' Croatian *liluška* 'passport', *liliška* 'letter,' *lil* 'passport' < Rom. *lil* 'letter, document' (Hugo 2009; Uhlik 1954: 21);
- Czech *raj* sec. 'master, judge,' Croatian *rain* 'policeman' < Rom. *raj* 'master, judge' (Hugo 2009; Kamiš 1998: 129; Uhlik 1954: 26);
- Czech *sovelit/suvelit* sec. prison sl., criminal sl. 'to sleep', *subelit* arg. 'ditto,' Croatian *soviti* 1. 'to sleep' 2. 'to break into, to follow' 3. 'to observe closely' < Rom. *sovel*, *sovav* 'I am sleeping, to sleep' (Hugo 2009; Kamiš 1998: 129; Uhlik 1954: 26–27).

3.2. Semantic categories of Romani borrowings

As mentioned earlier, words of Romani origin in the lexis of both languages are to a large extent connected with criminality and the administration of justice, which may stem from the former and the present social situation of the Roms in both countries. Moreover, this lexis is characterised by strong emotions, as seen in e.g. the category "names of persons." These are particularly frequent in the Czech vocabulary, being recorded 108 times, yet there are only in Croatian 41 such instances (cf. Table 1).

In the Czech material there are expressive designations for policemen as follows: *bengoš* criminal sl. 'policeman,' *benga/benka* criminal sl. 'policemen,' *bengálec* youth sl. 'policeman' < sec. *beng(o)* 1. 'gendarme' ← 'devil' < Rom. *beng* 1. 'devil' 2. 'evil, malicious child', 3. 'evil man' (Hugo 2009), *panglo* 'policeman' < *panglit/phanglit* prison sl. 'to confine, to imprison' < Rom. *phandel* 'to lash, to confine' (Hugo 2009), *šelengero/šilingero, šelengerák, šelengeres, šilingere* arg., criminal sl. 'gendarme,' *šilingr* 'policeman', *šilingrál* criminal sl. 'policeman' < Rom. *šelengero* 'rope-maker' < *šelo* 'string, rope' (Hugo 2009; Kamiš 1998: 129); designations of thieves: *čórák, čórář, čorkář, čór* criminal sl. 'thief' < Rom. *čor* 'ditto' (Hugo 2009), *kérař/khérař* criminal sl. 'burglar' < *kér/khér* 1. 'flat, property,' 2. arg. 'place of robbery' < Rom. *kher* 1. 'home' 2. 'flat' (Hugo 2009) as well as prisoners of various type: *čirykle* 1. prison sl. 'a weaker prisoner who serves a more powerful one' ← 2. sec. 'bird' < Rom. *čiriklo* 'bird' (Hugo 2009), *styldo* sec. 'prisoner'< Rom. *styldo* 'confined, imprisoned' (Hugo 2009). Expressive elements also characterise the names for poor men and beggars: *geróro* sec. 'poor man' < Rom. *gero* 'the deceased, dead man,' *mangelář* sec. 'beggar' < *mangelit* criminal sl. 1. 'to beg' 2. 'to ask' < Rom. *mangel* 'ditto' (Hugo 2009; Kamiš 1998: 129); designations for naive persons: *dyliňák* 'fool,' *dylina* 1. 'a foolish

---

[11] In the paper by Uhlik (1954: 20) there are also forms which are identical with the Czech ones: *kuliška, kul* 'excrement'.

woman' 2. arg. 'madman, fool' < Rom. *dylino/dilino* 'fool' (Hugo 2009), *máčovka* criminal sl. 'fool', *máčo* 1. prison sl. 'fool' 2. prison sl. 'a mentally ill man' < Rom. *máčo/máčho* 'fish' (Hugo 2009; Kamiš 1998: 129) and names for women: *čaje* sec. 1. 'girl' 2. 'woman' < Rom. *čhaj/čáj* 'Romani girl, daughter' (Hugo 2009; Kamiš 1998: 128), *jambora/gambora, jamborka* 1. sec. prison sl. 'woman, girl' 2. arg. 'woman who often changes partners' < Rom. *gamba* 'girl' (Kamiš 1998: 129).

In the Croatian material there are also numerous expressive designations for women; these are very often synecdoches of various parts of the body, usually the female genitals, e.g.: *hevina* 1. 'prostitute' ← 2. 'female genitals' < Rom *chiv, chuv* 'hole' (Uhlik 1954: 17), *mindžulja, mindža* 1. 'woman, girl' ← 2. 'female genitals' < Rom. *mindž/minž* 'female genitals' (Hugo 2009; Uhlik 1954: 24), also *dandara* 'a woman who talks a lot' < *danda* 'tooth' < Rom. *dand* 'ditto' (Uhlik 1954: 12–13). Besides, Romani provenience is also shown in the designations for naive persons as follows: *dileja* 1. 'fool' 2. 'insane' 3. 'naive', *dilkan* 1. 'mad' 2. 'foolish' 3. 'unpredictable' < *delina* 'blockhead' < Rom. *delino/dilino* 'ditto' < *dilo* 'mad' (Uhlik 1954: 13; Imami 2007) as well as for guards or law officers, e.g. *rain* 'policeman' < Rom. *raj* 'master, judge' (Uhlik 1954: 26).

Particularly frequent in the Croatian material are the "parts of the body" (30) category, which usually are of an intimate nature (some of them were mentioned in the previous paragraph), e.g.: *buljara, buljeskara* 'big bum,' *bulja* 1. 'head' 2. 'bum' < Rom. *bul* 'bum' (Uhlik 1954: 10; Imami 2007), *kandilo* 'bum' < *kandisati* 'stink' < Rom. *khandel* 'it stinks' (Uhlik 1954: 18), *kar, kardan, karson* 'male genitals' < Rom. *kar/ker* 'ditto' (Uhlik 1954: 19; Horbač 2006: 377), which in the Czech material appear 23 times, e.g.: *chynda* vulg. prison sl. 'bum' < Rom. *chindi* 'ditto' (Hugo 2009), *káro, kár* prison sl. 'male genitals' < Rom. *kár* 'ditto' (Hugo 2009; Kamiš 1998: 129), *muj* prison sl. 'mouth' < Rom. *muj* 'ditto' (Hugo 2009; Kamiš 1998: 129).

In both languages numerous designations for "human activities" can be found (Czech 16, Croatian 24), and mostly these are illegal activities, e.g. Czech: *čórka* criminal sl. 'theft,' *čorovačka* 'ditto' < *čórovat* sec. 'steal,' *mangelka/mengelka* 'beggary, begging' < *mangelit* criminal sl. 1. 'to beg' 2. 'to ask'; Croatian: *kidavela, kidanje* 'to run away' < *kidati* 'to run away' < Rom. *kidav* 'I take / I am taking, I run away / I am running away' (Uhlik 1954: 20), *marinela* 'theft,' *maravela, marisana* 'fight' < *marisati* 1. 'to steal' 2. 'to beat' < Rom. *mariv* 'I will kill/I will hit' (Uhlik 1954: 22–23) as well as for "places and institutions" (Czech 21, Croatian 12), which are usually directly linked with the illegal activities, e.g.: Czech: *čorokér/čorokhér* sec. arg. 'penitentiary, prison' < Rom. *čor* 'thief' + *kher* 'home' < Rom. *kher* 1. 'home' 2. 'flat' (Hugo 2009), *stylipen* 'detention'; Croatian: *kerna* 1. 'home' 2. 'stable, cowshed' 3. 'hiding place, shelter' 4. 'hotel' < Rom. *kher* 'home' (Uhlik 1954: 20), *kulana* 'prison'.

It is unsurprising to find alternative designates for "names of money" (Czech 12, Croatian 12) in both sets of material, most of which are various derivatives from still now popular Romani base *lóve*, e.g. Czech: *lovasy, lováče, lováky, love/lóve*[12]; Croatian: *lovica, lovijana, lovijanović, loviška, lovuta, lovušina, lovuša, lova, vula, valo*. There are also a similar number of designations for "names of clothes" (Czech 6), e.g.: *gad/*

---

[12]   But also the form *mari* arg. 'money' < Rom. *mari* 'ditto' (Hugo 2009; Kamiš 1998: 128).

gat[13] sec. 'shirt' < Rom. *gad* 'shirt' (Hugo 2009); (Croatian 7), e.g.: *giljka* 'shoe', *giljarica, giljarka* 1. 'shoe' 2. 'leg', *giljara, gilja* 1. 'shoe' 2. 'leg' 3. 'escape' < Rom. *geljom* 'I was walking' (Uhlik 1954: 16).

Moreover, in the Czech material there are numerous designations for "animals" (Czech 15, Croatian 3), e.g. Czech: *graj* sec. 'horse, pony' < Rom. *graj* 'horse' (Hugo 2009; Kamiš 1998: 129), *kahně* argot 'hen' < Rom. *káhni* 'ditto' (Hugo 2009; Kamiš 1998: 129), *murga* arg. 'female cat' < Rom. *murka* 'ditto' (Kamiš 1998: 129), Croatian: *đukac* 'dog', *đukela* 'dog (large)'; for "names of tools" (Czech 12, Croatian 1), usually linked with committing crimes, e.g. Czech: *čúro* argot prison sl. sec. 'knife' < Rom. *čúri* 'knife' (Hugo 2009; Kamiš 1998: 129), *karibengri/kharibengri* sec. 'revolver' < Rom. *karibengri* 'ditto' (Hugo 2009); for names of "food" (Czech 10, Croatian 3), e.g. Czech *máro* prison sl. 'bread' < Rom. *máro* 'ditto' (Hugo 2009; Kamiš 1998: 129), Croatian *hališka* 'food' < *halisati* 'eat' and for "stimulants" (Czech 7, Croatian 3), e.g. Czech *melardo* sec. 'coffee' < Rom. *melardo* 'soiled' (Hugo 2009), Croatian *kerija* 'rakia' < Rom. *kerki* 'bitter' (Uhlik 1954: 20).

Among the remaining categories further Czech examples include: *dárel* sec. prison sl. 'fear' < Rom. *daral* 'to fear' (Hugo 2009), *kúlo/khúlo* sec. prison sl. 'excrements, dung,' *rat* arg. 'night' < Rom. *rat* 'ditto' (Hugo 2009; Kamiš 1998: 129) as well as Croatian: *ćorka* 1. 'sleeping, sleep' 2. 'relaxation,' *duja* 1. 'tram line no. 2' 2. 'a failing mark, a school mark' < Rom. *duj* 'two' (Hugo 2009), *lovostaj* 'financial situation'.

As has already been mentioned above, in the analysed material there are also a large group of verbs of Romani origin – 87 Czech and 46 Croatian. To a large extent these are connected with various illegal activities or the criminal world, e.g. Czech: *bešelit* 1. sec. 'to sit' 2. prison sl. 'to be imprisoned' < Rom. *bešel* 'to sit' (Hugo 2009), *denášelit* sec. 'to run away' < Rom. *denašel* 'ditto' (Hugo 2009), *chudelit* 1. svět. 'to take, to catch' 2. prison sl. 'to steal, to rob' < Rom. *chudel* 'to take, to catch' (Hugo 2009), *našavelit* sec. prison sl. 'to kill' < Rom. *našavel* 'to lose' (Hugo 2009), *panglit/phanglit* sec. prison sl. 'to lock away, to imprison' < Rom. *phandel* 'to tie, to lock away' (Hugo 2009; Kamiš 1998: 129); Croatian: *delisati* 1. 'to deal cards in a card game' 2. 'to give something from among the stolen goods' < Rom. *dav, del* '(s/he) gives' (Uhlik 1954: 14), *hapati* 1. 'to steal' 2. 'to catch' 3. 'to eat' < Rom. *chape* 'food' (Uhlik 1954: 17), *marisati* 1. 'to steal' 2. 'to beat,' *marisati se* 'to fight' < Rom. *mariv* 'I will hit, I will kill' (Uhlik 1954: 22). The remaining verbs, although emotionally coloured, already have a more general meaning, e.g. Czech *ačelit/hačelit* sec. 'to remain, to sit for a while' < Rom. *ačel* 'to remain' (Hugo 2009), *dykchelit* sec. prison sl. 'to look' < Rom. *dikchel* 'ditto' (Hugo 2009; Kamiš 1998: 129), *džanelit* sec. prison sl. 'to know' < Rom. *džanel* 'to know' (Hugo 2009; Kamiš 1998: 129), *kérovat/khérovat* 1. sec. 'to do, to work' 2. sec. 'to haste' 3. sec. 'to be able to' 4. prison sl. 'to tattoo' 5. Brno sl. 'to speak, to persuade' 6. 'to punish someone' < Rom. *kerel* 'to work, to do' (Hugo 2009); Croatian: *đanisati* 1. 'to understand' 2. 'to look' 3. 'to sympathise with' 4. 'to like' < Rom. *džanav (džanam)* 'I know' (Uhlik 1954: 15), *penisti* 'to speak, to tell' < Rom. *phenav* 'I speak, I am speaking' (Uhlik 1954: 25).

---

[13]   This is the third word of Romani orgin which appears in the dictionary by Stępniak (1993).

Moreover, in the Czech material there are numerous adjectives, both negatively marked, e.g.: *džungalovný* sec. 'evil' < Rom. *džungalo* 'evil, ugly, dirty' (Hugo 2009), *kašuko* 1. sec. 'mute' 2. prison sl. 'blind' < Rom. *kašuko* 'deaf' – a shift of meaning (Hugo 2009), and positively marked, e.g.: *čáčo, čáčovný* sec. 'good' < Rom. *čačo* 'true, righteous' (Hugo 2009), *láčo, láčovný* prison sl. 'good' < Rom. *lačo* 'ditto' (Hugo 2009).

Semantic category	Number of lexemes	
	The Czech language	The Croatian language
Names of persons	108	41
Parts of the body	23	30
Places and institutions	21	12
Human activities	16	24
Animals	15	3
Money and shopping	12	12
Tools and equipment	12	1
Food	10	3
Stimulants	7	3
Clothing	6	7
Mental and physical states of the human being	5	2
Situations and events taking place in the community	5	1
Substances and materials	4	2
Furniture and furnishings	3	0
Vehicles	2	1
Measures	2	0
Music	2	0
Study	0	1
Other	7	4
Total:	**260**	**147**

Table 1. Borrowings from the Romani language – semantic categories of nouns

3.3. The morphological and word-formative structure of Romani borrowings

In both tongues, but especially in the Croatian material, the borrowed Romani forms in general adjust their morphological and word-formative form to the borrowing language. This may happen by means of a usual change or the addition of a native

grammatical morpheme, e.g. Czech: *čaje* sec. 1. 'girl' 2. 'woman' < Rom. *čhaj*, *chynda* vulg. prison sl. 1. 'jokes' 2. 'bum' < Rom. *chindi* 'bum'; Croatian: *dasa* 1. 'boy' 2. 'lover' 3. 'lady-killer' 4. 'dandy' < Rom. *Das* 'Serb, Croatian, non-Rom' (Uhlik 1954: 13), *đuva* 'an older person who has a lot of money' < Rom. *džuv* 'louse' (Uhlik 1954: 15). This is a common phenomenon among the borrowed verbs, e.g. Czech: *bikinelit* sec. prison sl. 'sell' < Rom. *bikinel* 'ditto' (Hugo 2009), *čórovat* sec. 'to steal' < Rom. *čorel* 'ditto'; Croatian: *delisati* 1. 'to deal cards in a card game' 2. 'to give something from among stolen goods' < Rom. *dav, del* '(s/he) gives, she is giving', *soviti* 1. 'to sleep' 2. 'to burgle, to follow' 3. 'to watch carefully' < Rom. *sovav* 'I sleep/I am sleeping'. In the case of nouns native suffixal morphemes are often added to the word,[14] e.g. Czech: *čajka* youth sl. 'girl' < Rom. *čaji* 'ditto', *darák* sec. prison sl. 'fear' < Rom. *dar* 'ditto' (Hugo 2009), *džúvák, džúvalák* sec. 1. 'beggar' 2. 'gendarme' < Rom. *džuv* 'louse', *džuvalo* 'lousy' (Hugo 2009; Kamiš 1998: 129), *minďoch* Brno sl. vulg. 'female genitals' < Rom. *mindž* 'ditto,' *zuminka* arg. prison sl. 'soup' < Rom. *zumin* 'ditto' (Hugo 2009); Croatian: *dasulja* 'old woman' < Rom. *Das* 'Serb, Croat, non-Rom' (Uhlik 1954: 13), *hevina* 1. 'prostitute' 2. 'female genital' < Rom. *chiv, chuv* 'hole', *kulana* 'prison' < Rom. *khul(m)* 'excrement' (Uhlik 1954: 20).

Further proof of words of Romani orgin having taken root in both Slavic languages are the derivatives resulting from a combination of the Romani and the Slavic bases (more often, however, stemming from other languages), e.g. Czech: *bengaboys* youth sl. 'policemen' < *beng* 'policeman' (< Rom. *beng* 'devil') + English *boys, čoklbuřt/čoklvuřt* derogat. 'cheap sausage', Brno sl. 'a dry smoked pork sausage' < Rom. *džukel* 'dog' + Germ. *Wurst* 'sausage'; Croatian: *kar-bunar* 'female genitals' < Rom. *kar* 'penis' + *bunar* 'well', *mindžocur* 'menstruation' < Rom. *mindž/minž* 'female genital' + *curiti* 'ooze,' *mindžosprej* 'deodorant for intimate body parts' < Rom. *mindž* + Eng. *spray*. In both sets of materials the results of the contamination process can also be found, e.g. Czech *mindžoleta, mindžoletka* 'woman of bad reputation' < Rom. *mindž* + *amoleta, amoletka* 'girl, girlfriend'; Croatian *lovčanik* 'wallet' < Rom. *lóve* + *novčanik* 'wallet', however, among the Croatian examples it is possible to find a particular type of metathesis, e.g. *valo* 'money' < *lova* 'ditto.'

Despite the adaptation tendencies described above it is also possible to find numerous examples of lexemes in the Czech material which have taken grammatical morphemes characteristic of Romani, e.g. *báro* 'big' < Rom. *báro* 'ditto' (Hugo 2009), *karialo* sec. 'meat' < Rom. *karialo* 'ditto' (Hugo 2009), *melali* sec. prison sl. 'coffee' < Rom. *melal'i* 'ditto' (Hugo 2009), *šingálo* prison sl. 'caretaker' < Rom. *šingalo* 'gendarme, policeman' (Hugo 2009).

---

[14] Naturally, they may also be treated as derivatives already formed in both Slavic tongues. The considerable productivity of some forms of Romani origin can be testified to, among others, by the derivatives of the Romani base *love* as noted in 3.2.

4. Frequency

The focus of the analysis so far has been on all the words of Romani origin found in Czech and Croatian lexicographic studies. The objective of this section is to identify those Romani words which are found in the Internet corpora of both languages, and how to analyse to what extent they are present in multifarious texts of the 20th and 21st c. It needs to be stressed initially that this analysis could not be carried out in a similar manner for both languages because Czech possesses a much more extensive corpus than Croatian. Without doubt, the most serious shortcoming of the Croatian corpus, in the context of the present analysis, is the lack of data based on Croatian spoken language.

In order to study the frequency of the words of Romani origin in the Czech language 7 reference corpora were used, of which 4 were corpora of spoken language (BMK, PMK, ORAL 2006, ORAL 2008) and 3 were corpora of written language (SYN 2000, SYN 2005, SYN 2010). In the case of the Croatian material the analysis was limited to one extensive corpus of written language, in which journalistic texts predominated (HNKv2.0). The results of the study are presented in Table 2 (Czech material) and Table 3 (Croatian material); in both cases the starting point is the Romani source word which may take various forms and meanings in the Slavic languages (cf. section 3.3). In the investigation into the frequency of Czech Gypsy, words found in the written language corpora which were elements of a longer utterance in the Romani language, mostly fragments of dialogue in belles lettres texts, were not included.

Among the 163 Romani etymons from which the Czech Gypsy words have been formed, continuations of 25 source words were found in the corpus material, i.e. 15% of the total. It is worth noting that none of the etymons is represented simultaneously in all the 7 data bases – the most widespread are continuations of the words *džukel* and *khulo*, which are not only recorded in the Prague corpus of spoken language (PMK). All the forms of Romani provenience are found in total in the corpus material 911 times (the most frequent are lexemes derived from the forms *gádžo* – 250 and *džukel* – 233). It needs to be emphasised, however, that unquestionably the majority of the Czech Romani borrowings are to be found in written language data bases (862 attestations, most of them in SYN 2005 – 339), whereas in the spoken language corpora only 49 are noted, that is a little over 5%. This is very important as between 15% (SYN 2000) and 40% of the texts (SYN 2005, SYN 2010) included in the written corpora are works of fiction whose authors (e.g. J. Topol, V. Třešňák) introduce words from the old argot (often of Romani origin) with the aim of achieving a certain archaisation or poetisation of the language. Such examples of Gypsy vocabulary should be treated as features of the writer's idiolect and not as a reflection of the actual frequency of the words in question.[15]

---

[15] On the other hand, in the corpora mentioned above between 33% (SYN 2005, 2010) and 60% (SYN 2000) of the texts are of a journalistic nature which provide more reliable information about the popularity of the given forms. Possibly in a further analysis two corpora of journalistic texts could be used: SYN 2006 PUB and SYN 2009 PUB.

Romani etymon	BMK	PMK	ORAL 2006	ORAL 2008	SYN 2000	SYN 2005	SYN 2010	SUM
ačel 'to remain'	–	–	–	–	–	2 *hačej, hačejte* 'ditto'	1 *hačej*[16]	3
beng 1. 'devil' 2. 'evil, malicious child' 3. 'evil man'	–	–	–	2 *benga* 'policemen'	–	–	3 *benga*	5
čhinel 'to write, to cut, to carve'	–	2 *čimelík* 'servant'	–	–	–	3 *čimelík*	–	5
čor 'thief', *čori* 'theft', *čorel* 'to steal, to rob'	–	1 *čórnout* 'to steal'	2 *čórkař* 'thief'	–	11 (4 × *čóra* 'theft', 3 × *čórka* 'theft', 2 × *čórkař*, 1 × *čóro* 'theft', 1 × *čórkařský* 'thievish')	34 (10 × *čórnout*, 7 × *čóro*, 7 × *čórkař* / *čorkař*, 3 × *čórka*, 2 × *čórovat* 'to steal', 2 × *čór* 'theft', 1 × *čórovaný* 'stolen', 1 × *čórlý*, 1 × *čórnuty* 'stolen')	11 (4 × *čórnout*, 3 × *čórovat* / *čorovat*, 2 × *čórkař*, 1 × *čóro*, 1 × *čorka*)	59

[16] If no meaning is provided in the inverted commas next to the Czech forms, this means that their meaning is the same as the identical forms in the older corpora.

Romani etymon	BMK	PMK	ORAL 2006	ORAL 2008	SYN 2000	SYN 2005	SYN 2010	SUM
degeša 'rabble, riff-raff' *degeš* 'dirty'	–	–	–	–	–	4 *degešák* 'a despicable man'	–	4
dylino (dilino) 'blockhead'	–	–	1 *ta dilinos* 'a derogatory term for a woman'	–	13 (12 × *dylina / dilina* 'a derogatory term for a man or a woman', 1 × *dyliňák* 'ditto')	4 *dylina*	11 *dylina*	29
džukel 'dog'	1 *čokl* 'ditto'	–	6 *čokl*	5 *čokl*	35 (33 × *čokl*, 1 × *žukl* 'ditto', 1 × *čoklíček* 'ditto')	85 (82 × *čokl*, 3 × *čoklík* 'ditto')	101 (94 × *čokl*, 7 × *čoklík*)	233
gádžo 'a white man, a non-Rom, a peasant'	–	–	–	–	118 (105 × *gádžo / gadžo* 'ditto', 9 × *gádžovský / gadžovský* 'concerning non-Roms,' 1 × *gadžovský* 1 × *gádžovka* 3 × *gádžovka/ gadžovka* 'a non-Rom woman')	71 (60 × *gádžo / gadžo* 7 × *gádžovský / gadžovský, gadžovský* 4 × *gádžovka / gadžovka*)	61 (52 × *gádžo / gadžo,* 7 × *gádžovský,* 1 × *gádžovka,* 1 × *gádžův* 'belonging to a non–Rom')	250

Romani etymon	BMK	PMK	ORAL 2006	ORAL 2008	SYN 2000	SYN 2005	SYN 2010	SUM
graj 'horse'	–	–	–	–	–	23 *graj* 'ditto'	–	23
chal 'to eat', *chaben* 'food'	1 *chálka* 'ditto'	–	–	–	–	1 *chalka*	–	2
kher 1. 'home' 2. 'flat'	–	–	–	–	2 *kér* 'ditto'	–	–	2
kerel 'to work, to do'	–	–	1 *kérovat* 'to speak'	4 (2 × *pokérovaný* 'tattooed', 1 × *kérko* 'tattoo', 1 × *kérovat* 'to punish')	7 (4 × *kérko* 3 × *pokérovaný*)	12 (7 × *pokérovaný* 2 × *kérovaný* 'tattooed' 1 × *vykérovat* 'to tattoo' 1 × *zkérovat* 'to arrange, to organise' 1 × *khérko* 'tattoo')	23 (12 × *kérko* 4 × *pokérovaný* 2 × *kérovat* 1 × *kérkař* 1 × *zkérovat* 'tattooed' 1 × *zkérovaný* 'tattooed' 1 × *pokérovat* 'to cover with writing, to scribble' 1 × *vykérovat*)	47

Romani etymon	BMK	PMK	ORAL 2006	ORAL 2008	SYN 2000	SYN 2005	SYN 2010	SUM
khulo 'excrements, dung'	3 (2 × *kulový* 'dung, shit, nothing, all the same, refusal', 1 × *kulovka* 'ditto')	–	5 *kulový*	4 *kulový*	12 (11 × *kulový*, 1 × *kulovka*)	37 *kulový*	34 *kulový*	95
lohe 'stolen goods'	–	–	–	–	1 *lohnout* 'to steal, to deceive'	9 *lohnout*	6 *lohnout*	16
lóve/love 'money'	7 *love* 'ditto'	–	3 *love*	–	13 (11 × *love / lóve*, 2 × *lováče* 'ditto')	8 *love / lóve*	1 *love*	32
máčo 'fish' (prison sl. 'blockhead')	–	–	1 *makovec* 'blockhead'	–	3 (1 × *máčové* 'ditto', 1 × *mák* 'ditto', 1 × *makovej* 'ditto')	8 (6 × *makovec*, 1 × *máčo* 1 × *mák*)	3 (1 × *máčo*, 1 × *makový*, 1 × *makovec*)	15
mangel 'to beg'	–	–	–	–	3 (1 × *mangelit* 'ditto', 1 × *mangelení* 'begging', 1 × *vymangelit* 'to elicit by asking')	1 *vymangelit*	–	4

Romani etymon	BMK	PMK	ORAL 2006	ORAL 2008	SYN 2000	SYN 2005	SYN 2010	SUM
mari 'money'	–	–	–	–	7 (4 × *mařeny* 'ditto', 3 × *máry* 'ditto')	21 (18 × *mařeny*, 3 × *máry*)	4 (3 × *mařeny*, 1 × *mařenky* ditto')	32
mindž 'female genital'	–	–	–	–	4 *minda* '2 × female she-cat, 1 × vagina 1 × woman'	10 (9 × *minda*: '7 × she-cat, 2 × woman', 1 × *mindža* 'woman')	2 (1 × *minda* 'vagina', 1 × *mindža* 'ditto')	16
more 'mate, man!'	–	–	–	–	2 *more* 'ditto'	–	–	2
mulo 'dead'	–	–	–	–	1 *multna* 'cell'	–	–	1
raj 'master'	–	–	–	–	–	1 *raj* 'ditto'	–	1
šelengero 'rope-maker', *šil* 'winter, chill'	–	–	–	–	10 (9 × *šilcárna* 'police station', 1 × *šilc* 'policeman')	1 *šilcárna*	–	11

Romani etymon	BMK	PMK	ORAL 2006	ORAL 2008	SYN 2000	SYN 2005	SYN 2010	SUM
táto 'warm'	–	–	–	–	–	1 *tátoš* 'homosexual'	–	1
ventra 'louse'	–	–	–	–	19 *ventra* 'derogatory term for a human, especially for a woman'	3 *ventra*	1 *ventra* 'derogatory term'	23
TOTAL	**12**	**3**	**19**	**15**	**261**	**339**	**262**	**911**

Table 2. Romani borrowings in the Czech language – frequency

Complementing the above data, the results of a frequency analysis of the Gypsy words found in the pre-war monograph by František Oberpfalcer (1934) are also worth mentioning, since apart from the lexemes which are identical with those found in the *Slovník nespisovnej češtiny*, half of the entries include different phonetic realisations, additional derivatives as well as different meanings of the words discussed. An analysis of 34 lexemes derived from 24 Romani etymons showed the absence of these archaic forms in online Czech corpora.

In the Croatian corpus of written language one may find continuations of 11 Romani etymons, which constitutes ca. 26% of all the analysed Romani source words. Romani words appear in the corpus material 548 times, out of which as many as 499 lexemes – i.e. more than 90% – are forms derived from the word *lóve/love*, which confirms the unparalleled popularity of the lexeme *lova* 'money' in colloquial language, in a broad sense of the word. It is also very productive, which can be corroborated by the presence in the database of some derivatives formed from this base: *lovaš, lovator, lovatorica, lovica*. Additionally, there are 29 attestations in the corpus of the forms *kidati, kidnuti*, in the sense of 'to run away,' although their Romani provenience, appears to be debatable.

Romani etymon	HNK_v2.0 Croatian forms	Number
čor 'thief'	3 × *ćorka* 'prison', 1 × *ćorkirati* 'to imprison'	4
delino (*dilino*) 'fool' *dilo* 'insane'	1 × *dileja* 'blockhead'	1
duj 'two'	1 × *duja* 'a school mark'	1
džukel 'dog'	2 × *džukela* (1 × 'dog', 1 × 'derogatory term for a woman'), 1 × *džukac* 'dog'	3
geljom 'I was walking'	4 × *giljati* 'to run, to fight, to toil'	4
chape 'food'	1 × *hapanje* 'theft', 1 × *hapati* 'to steal'	2
kar (*ker*) 'male genital'	2 × *karati* 'to have sexual intercourse'	2
kidav 'I take / I am taking, I run away / I am running away'	21 × *kidati* 'to run away', 8 × *kidnuti* 'to escape'	29
lóve / love 'money'	487 × *lova* 'money', 8 × *lovaš* 'rich man', 2 × *lovator* 'rich man', 1 × *lovatorica* 'rich woman', 1 × *lovica* 'money'	499
mari 'money'	2 × *marijaš* 'ditto'	2
mariv 'I will hit, I will kill'	1 × *izmarati* 'to beat up, to massacre'	1
TOTAL:		**548**

Table 3. Romani borrowings in Croatian – frequency

5. Conclusions

The comparative analysis of the Czech and Croatian lexis of Romani origin has confirmed the thesis initially postulated that this type of vocabulary would be similar in both of the languages studied. This is attested by the presence of about 20 Czech-Croatian word pairs with the same Romani etymon, the same meaning and a similar form, as well as by a considerable number of lexemes which, despite a certain semantic difference, are linked by a common source word. Moreover, the same semantic categories dominate in both lexical materials, namely the strongly emotionally coloured "names of persons," "parts of the body" (usually connected with the intimate zone), illegal "human activities" as well as "names of money." In terms of adaptation, the Czech and Croatian Gypsy words in the main adjust their morphological and word-formative form to the borrowing language.

The investigation into the frequency of the words of Romani origin demonstrated that continuations formed were from 15% of the Romani etymons present in the Czech corpora which were the focus of the analysis, yet in the Croatian corpus, there were continuations in 26% of the source words. The longevity of some Romani borrowings is also attested by the fact that in the Czech corpus material forms of Romani provenience are found 911 times, but in the Croatian corpus there are only 548 examples. These studies are certainly not totally conclusive as in the case of the Croatian material they only involved written language. Thus it would seem expedient to conduct a questionnaire study to provide more complete results regarding the similarity between forms of Romani provenience among users of the Czech and Croatian languages.

It also seems necessary to explain the reasons behind the use of words with the same Romani etymon that appear in the lexicon of both languages as well as their vestigial presence in the lexis of the Polish language. Thus it might also be worth analysing other Slavic languages to further our knowledge on the subject. This, however, would require further interdisciplinary research providing insights into the specific influence of the language and culture of the Roms – seen as a group existing in social isolation – on the language and customs of Slavic nations. Without any doubt, at this point we might consult extensive literature dealing with various borrowing mechanisms. An attempt at answering the question why Romani words have not been adopted even by the Polish criminal underworld would certainly be a challenging task.

References

1. Source materials and dictionaries

BMK = Český národní korpus BMK. Praha 2002. [avaiable at: http://www.korpus.cz].
HNKv2.0 = *Hrvatski nacionalni korpus – HNKv2.0.* Zagreb 2005. [avaiable at: www.hnk.ffzg.hr].
ORAL 2006 = *Český národní korpus – ORAL2006.* Praha 2006. [avaiable at: http://www.korpus.cz].

ORAL 2008 = *Český národní korpus – ORAL2008*. Praha 2008. [avaiable at: http://www.korpus.cz].

PMK = *Český národní korpus – PMK*. Praha 2001. [avaiable at: http://www.korpus.cz].

SYN 2000 = *Český národní korpus – SYN2000*. Praha 2000. [avaiable at: http://www.korpus.cz].

SYN 2005 = *Český národní korpus – SYN2005*. Praha 2005. [avaiable at: http://www.korpus.cz].

SYN 2010 = *Český národní korpus – SYN2010*. Praha 2010. [avaiable at: http://www.korpus.cz].

Anić V., Goldstein I. 2000. *Rječnik stranih riječi*. [2nd edition] Zagreb.

Holub J., Lyer S. 1968. *Stručný etymologický slovník jazyka českého*. Praha.

Hugo J. 2009. *Slovník nespisovné češtiny*. [3rd edition] Praha.

Imami P. 2007. *Beogradski frajerski rečnik*. [3rd edition] Beograd.

Klaić B. 2001. *Rječnik stranih riječi*. Zagreb.

Rejzek J. 2001. *Český etymologický slovník*. Voznice.

Sabljak T. 1981. *Šatra (rječnik šatrovačkog govora)*. Zagreb.

Sabljak T. 2001. *Rječnik hrvatskoga žargona*. Zagreb.

Skok P. 1971–1974. *Etimologijski rječnik hrvatskoga ili srpskoga jezika*. [vol. 1–4]. Zagreb.

Stępniak K. 1993. *Słownik tajemnych gwar przestępczych*. London.

2. Monographs and articles

Buttler D. 1978. Kategorie semantyczne leksyki potocznej. – Szymczak M. (ed.). *Z zagadnień słownictwa współczesnego języka polskiego*. Wrocław: 37–45.

Chloupek J. 2003. Diferenciace a stratifikace národního jazyka. – Čechová M. (ed.). *Současná česká stylistika*. Praha: 45–58.

Fink Ž. 2003. Šatrovački govor u Hrvatskoj jučer i danas. – *Wiener Slawistischer Almanach* 52: 79–90.

Horbač O. 2006. *Argo v Ukrajini*. Ľviv.

Grabias S. 2001. Środowiskowe i zawodowe odmiany języka – socjolekty. – Bartmiński J. (ed.). *Współczesny język polski*. Lublin: 235–253.

Kamiš K. 1998. Několik poznámek k romským argotismům v mluvě české majority. – Klimeš L. (ed.). *Sborník přednášek z VI. konference o slangu a argotu v Plzni 1998*. Plzeň: 126–131.

Krčmová M. 1997. Současná běžná mluva v českých zemích. – Daneš F. (ed.). *Český jazyk na přelomu tisícletí*. Praha: 160–172.

Oberpfalcer F. 1934. Argot a slangy. – V. Dedina (ed.). *Československa vlastivěda III. Jazyk*. Praha: 311–376.

Oczkowa B. 2005. Obecność normy w świadomości użytkowników języka, czyli jak mówią Chorwaci. – Dąbrowska-Partyka M. (ed.). *W poszukiwaniu nowego kanonu. Reinterpretacje tradycji kuturalnej w krajach postjugosłowiańskich po 1995 roku*. Kraków: 85–96.

Šehović A. 2009. Leksika razgovornoga bosanskog jezika. – Halilović S., Tanović I., Šehović A. (eds.). *Govor grada Sarajeva i razgovorni bosanski jezik*. Sarajevo: 111–199.

Sgall P., Hronek J. 1992. *Čeština bez příkras*. Praha.

Silić J. 1997. Razgovorni stil hrvatskoga standardnog jezika. – *Kolo* 6. 4: 483–495.

Uhlik R. 1954. Ciganizmi u šatrovačkom argou i u sličnim govorima. – *Glasnik Zemaljskog muzeja u Sarajevu* 9: 5–31.