Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech

Inga Vyšniauskienė

Research Institute of the Lithuanian Language

inga.dervinyte@flf.vu.lt, idervinyte@gmail.com

Abstract

This paper presents the quantitative analysis of Russian and English resources in Vilnius adolescents' language. The analysis draws on Vilnius Adolescents' Language Corpus data, documenting the linguistic practices of 10–16 year old adolescents from differently socially and ethnically marked Vilnius neighbourhoods. The research has revealed strong correlation between ethnic marking of Vilnius neighbourhoods and both frequency as well as lexico-functional variation of Russian resources. Creativity, playfulness, situational strategic use of Russian resources characterize the language of adolescents from ethnically marked neighbourhoods. In contrast, Russian resources among adolescents in ethnically unmarked neighbourhoods have been limited to slang and swearing. In addition, Russian resources may correlate with such values as toughness and masculinity as they were twice more frequent among boys (from ethnically unmarked neighbourhoods) than girls. English resources are available to adolescents through computer-mediated communication, popular culture. Accordingly, the frequency of English resources and lexico-functional categories are similar among adolescents irrespective of the ethnic marking of the district. Generally, Vilnius adolescents' practices of using Russian and English resources can hardly be described as one system. It is rather an open, dynamic interplay of different sub-systems having their internal organization, norms, which correlate with gender or ethnic marking of Vilnius neighbourhoods.

Key-words: sociolinguistics, Vilnius speech, adolescents' language, Vilnius adolescents' language corpus, English resources, Russian resources.

Raktažodžiai: sociolingvistika, Vilniaus kalba, paauglių kalba, Vilniaus paauglių kalbos tekstynas, angliški ištekliai, rusiški ištekliai.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

2

1. Introduction

This paper presents the corpus analysis of Vilnius adolescents' heteroglossic practices[1]. Adolescents' linguistically diverse practices have been relatively well analysed in major European cities (Blommaert and Rampton 2011; Forsskähl 2001; Leppänen 2007) but little yet explored in Lithuania (Čekuolytė 2012; Vyšniauskienė 2012). Sociocultural, demographic and ethnic diversity of major cities provides various linguistic resources for adolescents, which they use unpredictably, creatively and meaningfully. It is, therefore, not a coincidence that adolescents' linguistic practices have been fruitfully analysed in urban spaces. The capital city of Lithuania, Vilnius has been chosen as a fruitful niche for this research as it is Lithuania's most culturally and linguistically diverse city with residents displaying highest levels of language knowledge, including the languages of ethnic minorities, namely Russian and Polish, as well as other foreign languages, with English ranked highest (Nevinskaitė 2010: 56).

Obviously, Vilnius adolescents' linguistic practices do not exist in a vacuum but are closely associated with particular values attributed to Russian and English[2], which may be different in society at large as well as among adolescents.

As shown by attitude surveys, the functions and value ascriptions of Russian and English are not equally distributed (Vaicekauskienė 2010: 175). This relates to historical and social changes of the role of Russian and English. English has experienced a particular growth in value since the reestablishment of the Lithuanian state in 1990 (Vaicekauskienė 2014). Namely, attitude survey of Vilnius city dwellers has shown that English carries overt prestige, it is associated with proficiency and instrumental value, positive allegiance with the West (Vaicekauskienė 2010: 194; cf. Verschik on similar evaluation of English in Estonia (2010)). Russian, on the other hand, has been downgraded in value since the regained independence, but has retained its communicative function. Russian is mostly used for communication with ethnic minorities as 20 % of Vilnius dwellers have Russian as their native language (Vaicekauskienė 2014). 15–24 year old youth claims to have the best command of English,

---

[1] The present research was carried out during the research project "Vilnius Speaking II: city and urban linguistic diversity", carried out by the Research Institute of the Lithuanian Language, funded by the Research Council of Lithuania project leader Dr., Assoc. Prof. Loreta Vaicekauskienė, VAT–61/2012.

[2] The values associated with Polish are not discussed as only few Polish resources have occurred in the corpus data.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

3

whereas young speakers' ability to use Russian is lower in comparison to older age groups (Vaicekauskienė 2010).

Indexical values of Russian and English are negotiable, susceptible to changes or re-valuation. Research of Vilnius adolescents' linguistic practices has shown that Russian and English resources acquire different values in adolescents' speech (Čekuolytė 2012). Russian words and swearwords are mostly used by "tough" boys, since such language allows demonstrating power, physical strength and masculinity (ibid.). This is in line with other pieces of research, which have shown that boys need to use strong, shocking expletives as a symbol of masculinity (Andoroutsopoulos and Geograpoulou 2003: 109). English quotations are more frequent among girls interested in pop-culture (Čekuolytė 2012). In addition, English resources tend to be valued as indexing modernity, fun, and smartness (ibid.) (cf. similar value ascriptions of English resources in Leppänen (2007)).

Research of Vilnius adolescents' speech, although sporadic (Čekuolytė 2012; Vyšniauskienė 2012), has shown that adolescents have developed a systematic organization of linguistic resources they use. This system and values may not coincide with the values of Russian and English in society at large but they cannot be described as the inadequate mixture, unnecessary insertion of Russian and English words, as commonly voiced in the public discourse on youth speech. The present corpus analysis of Vilnius adolescents' speech serves as a continuation of previous, although significantly smaller in scope, pieces of research (ibid.). The paper is intended as a scientific argument for the existence of norms for using Russian and English resources among adolescents. Specifically, the present research draws on Vilnius Adolescents' Language Corpus[3] data and aims to identify what norms exist with reference Russian and English resources along the parameters of gender and ethnic marking of Vilnius neighbourhoods.

Theoretically, Vilnius adolescents' linguistic practices are approached post-structurally, conceptually rather than structurally. Focus is laid on the unpredictable, creative and meaningful use of resources available to adolescents rather than on describing the limits, structure and number of languages employed by adolescents.

---

[3] The corpus was collected during the during the research project "Vilnius Speaking II: city and urban linguistic diversity" (see Footnote 1).

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

4

In Section 2 I will provide a brief overview of the conceptual challenges that a quantitative analysis of adolescents' language may raise as well as expand on the notion of resources. Section 3 will present Vilnius Adolescents' Language Corpus composition and its collection criteria. The analysis of the corpus data is presented in Section 4.

2. Languages or resources: in search of an analytical category

Research of adolescents' linguistic practices in urban spaces has shown that the identification of the source language of slang and swearwords can indeed be problematic. First, pre-given category labels like Russian or English cannot be applied unless they reveal participant perspective. Cáccamo (2002: 29) distinguishes between the linguistic material of utterances, traditionally termed as linguistic varieties, and communicative codes, i.e. the associative mechanisms which underlie their production. As exemplified in various pieces of research, interacting in what are traditionally thought of as distinct languages, may in fact be one mixed-code for speakers (Cáccamo 2002: 30; Blommaert 2005: 76; Otsuji and Pennycook 2009: 241). Similarly, Blommaert (2013: 614) places emphasis on the perpetual processes of creative coding, the continuous production of new codes, which are not tied to the traditional notion of 'languages'. Consequently, the present research should reveal the categories that are relevant to the participants of the conversation, and they need not necessarily coincide with distinct languages in the classical sense (see Section 3.3).

The difficulty related to the analysis of Vilnius adolescents' linguistically diverse repertoire is also associated with the notion of a language. In other words, linguistic diversity observed in adolescents' speech seems to raise analytic challenges to the traditional approaches based on the notion of 'a language'. Scholars claim that the nondynamic concept of 'languages' as bounded entities cannot be adequate for analyzing linguistic diversity of urban spaces (Makoni and Pennycook 2007; Blommaert 2013). Such debates have resulted in varied terminology in youth language research, among which 'crossing' (Rampton 2010) and 'polylanguaging' (Jørgensen, Karrebaek and Madsen 2011) have received substantial scholarly attention. The framework of 'polylingualism' has been proposed to capture linguistic features that adolescents have at their disposal to achieve their communicative aims as best they can, e.g. young Danes using lexical, phonological features of Arabic, French, Punjabi. Although 'crossing' refers to a similar linguistic phenomenon, emphasis is laid on the use of linguistic

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

5

resources which are generally thought not to belong to the speaker, e.g. London teenagers using Punjabi, and, hence, implies crossing of social and ethnic boundaries. 'Polylingualism' and 'crossing' presented a conceptual turn from plurality or multiplicity perspective (as exemplified by 'multilingualism', 'bilingualism') to the notion of complexity of language (Blommaert 2013: 613). Adolescents are perceived as 'languagers' who purposefully use resources from languages they have little or no command of (Jørgensen, Karrebaek and Madsen 2011; Rampton 2010).

Youth language does not easily lend itself to quantitative analysis. Namely, even if the notion of a language is replaced with the notion of a set of features (Jørgensen, Karrebaek and Madsen 2011), the researcher's focus may fall on counting and enumeration rather than the significance of such a linguistic practice to participants. In addition, it may still be problematic to identify the boundaries of features as well as categorise them into languages. As Jørgensen , Karrebaek and Madsen (ibid.) claim, it is difficult to identify whether youth Danish, Danish with an accent would be treated as one category or separate categories. In other words, different sets of resources may form hybrid forms, in which case it is impossible to determine one set of resources as fundamental.

Moreover, as advocated by youth language researchers working within the framework of polylingualism and crossing, rather than separate languages, scholarly analysis should primarily focus on what such linguistic practice means to adolescents, thus claiming the benefits of interactional sociolinguistics. In addition, quantitative analyses on adolescents' linguistic practices may end up following a correlational fashion of relating frequency with identity. For example, Forsskähl (2001: 1) claims that geographical and gender identity can be conveyed through the choice of different slang words. Similarly, Lėgaudaitė (2005: 177) posits that frequent slang terms in boys' speech reveal masculine identity and peer-group (of a certain Kaunas neighbourhood) identity. However, as demonstated by Rampton (2010) on crossing as well as Jørgensen, Karrebaek and Madsen's (2011) research of polylanguaging youth, the connection between linguistic repertoire and identity is not absolute and straightforward. Working within the poststructuralist view of language and fluid nature of youth speech, analyses shift away from a focus on how distinct codes are switched or mixed to analyzing for what purpose/style construction linguistic resources are used. It should be mentioned that the social constructionist approach lays focus on the discursive construction of identities. Hence,

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

6

corpus analysis should not pose direct mapping between most frequent linguistic forms and gender identity (Bucholtz 2009: 146; Otsuji and Pennycook 2009: 241). A quantitative analysis cannot provide answers to questions on identity negotiation. Nevertheless, it can reveal valuable correlations between linguistic practice and groups of users for whom such linguistic practice is important. As Verschueren claims, "<…> investigations along quantifiable parameters may provide guidance to an assessment of their character and importance" (1991: 188).

The notion of a language is conceptually inadequate for analysing adolescents' linguistic practices. Accordingly, the present paper will apply another analytic level. Namely, the research will be based on Blommaert's (2010, 2013) social approach to language as a set of resources. The notion of resources encompasses dynamic aspects of linguistic practice, which would not be captured by the notion of a language. It is likely that the present corpus analysis will provide a more dynamic perspective on Vilnius adolescents' linguistic diversity than offered by the languages as separate systems ideology.

3. Research data

Vilnius Adolescents' Language Corpus (henceforth VALC) has been compiled for the present study. The presentation of VALC composition in the following Section is supplemented with information on its collection criteria as well as some practical obstacles. In Section 3.2 I overview the non-linguistic parameters of corpus collection, namely, ethnic and social marking of Vilnius neighbourhoods. In Section 3.3 I discuss the coding of Russian and English resources.

3.1 Vilnius Adolescents' Language Corpus: composition and data collection

Vilnius Adolescents' Language Corpus, which is partly modelled on the design criteria of the Bergen Corpus of London Teenage Language (COLT), the Corpus of Spanish Teenage Language[4], is the first large collection documenting 35 hours[5] (approximately 170 thousand words) of 10–16 year old Vilnius

---

[4] The Bergen Corpus of London Teenage Language is a collection of 50 hours of spontaneous (half a million words) speech of 13–17 year old teenagers from different neighbourhoods of London. The Corpus of Spanish Teenage Language/The Corpus Oral de Lenguaje Adolescente (COLA) has been modelled on the design criteria of COLT (see Stenström, Andersen and Hasund, 2002).

[5] To make the corpus comparative to other major corpora of teenage speech, it is planned to expand the corpus to 50 hours of Vilnius adolescents' spontaneous speech.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

7

adolescents' spontaneous speech in ethnically marked (henceforth EthM), i.e. mostly populated by ethnic Russians and Poles, and ethnically unmarked (henceforth EthUn), i.e. ethnic Lithuanians forming the majority, Vilnius neighbourhoods (see Table 1).

Table 1. VALC composition according to neighbourhoods' ethnic marking and gender: size, number of informants-volunteers and participating speakers

	EthM		EthUn		Total in both types of neighbourhoods	
	Boys	Girls	Boys	Girls	Boys	Girls
No of informants	8	13	25	31	33	44
Total	21		56		77	
No of participants	22	30	60	70	82	100
Total	52		130		182	
No of words	15,315	36,896	45,430	68,027	100,368	116,371
	52,211		113,457		165,668	

77 informants (21 attending schools in EthM neighbourhoods and 56 informants attending schools in EthUn neighbourhoods) were school students chosen randomly in collaboration with teachers, who often encouraged the students to take part in the project. Obviously, before approaching the informants, research ethics was considered, i.e. the headmaster's permission as well as parents' permission for their child to participate in the project had to be received.

Despite full anonymity guaranteed for all the participants, relaxed atmosphere, friendly researcher and good grades promised by some of the teachers, boys were less reluctant to act as informants in the project, which resulted in more female informants, i.e. 33 volunteering boys and 44 girls. The total number of participants has amounted to 182 as the recordings were made between the informants and their friends. One-gender conversations (boys talking to boys and girls interacting with girls) tended to predominate over mixed-gender conversations, which resulted in the total number of female speakers (100) slightly higher than the total number of male speakers (82). The minor difference in numbers did not appear significant as girls and boys contributed roughly the same number of words: boys about 46 per cent (100,363 words) and girls 54 per cent (116,371 words). Interestingly, similar proportions across gender categories are represented in COLT, with 48 per cent boys' and 52 per cent girls' speech (see Stenström, Andersen and Hasund 2002: 6).

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

8

The sub-corpus of adolescents' language from EthUn neighbourhoods is twice larger in comparison to the data from adolescents in EtnM neighbourhoods. The number of speakers, whose speech samples constitute the sub-corpora, also differs significantly: 52 speakers in EthM neighbourhoods and 130 participants in EthUn neighbourhoods. Several reasons can account for this difference. First, there are only few neighbourhoods in Vilnius which could be defined as EthM in comparison to five EthUn neighbourhoods chosen for the project (see Section 2.2.), hence, limiting the range of schools and informants. In addition, adolescents from EthM neighbourhoods appeared reluctant to their participation in the project, provided fewer recordings suitable for transcription in comparison to adolescents from EthUn neighbourhoods. Adolescents from EthM neighbourhoods were also often unconfident and susceptible that their language might be evaluated from the normative perspective.

School-students who volunteered to record their conversations with friends were instructed about the procedure, requirements and circumstances of the recording as well as introduced to research ethics and anonymity. The informants were equipped with a portable digital recorder and a lapel microphone (though not all the informants used it) for one week and instructed to record at least two hours of spontaneous speech with friends in their leisure time. In addition, the informants were asked to record in not very noise places without, if possible, their friends noticing the recorder. The informant had to tell friends about the recording at the end of the recording process and ask for friends' permission for their language having been recorded and be used for research purposes. If some adolescents, however, felt uncomfortable recording their friends, they could inform their friends about the recording and ask for their permission at the beginning of the recording process. When the major part of the corpus data was collected, it appeared that friends noticed the recorder at the beginning or in the middle of the recording process more often than not. Awareness of the recorder did not seem to make the participants less spontaneous but rather encouraged them to be bolder indulging into dirty jokes, or using swearwords when addressing the researcher (cf. similar difficulties when compiling COLT documented by Stenström, Andersen and Hasund (2002: 6)).

Informants were also asked to fill in the questionnaire form with information about the speaking situation (location, e.g. at home, school, canteen, café, etc.) and all the speakers, i.e. the informant (contact information, age) and other participating speakers (their real names, age, relation to the

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

9

informant, e.g. friend, classmate, sister). Obviously enough, the latter information remained confidential and was used for research purposes only.

The requirements overviewed were followed by the informants in the majority of cases. Most conversations, as asked, were recorded among peers. However, some informants' family members as well as teachers or strangers (shop assistants, neighbours) took part in several conversations. As the present analysis aims to reveal the distribution of Russian and English resources among 10–16 year old adolescents, the language of speakers who did not fall within the latter age category, was excluded from the present research. In addition, some speakers only came and left during the conversation. Accordingly, the speech samples of speakers who produced 100 words or fewer (7 girls and 4 boys) were excluded from the present analysis.

3.2 Non-linguistic parameters of VALC collection: ethnic and social marking of Vilnius neighbourhoods

For the corpus data to include adolescents' speech samples from Vilnius neighbourhoods of different ethnic and linguistic marking, a contrastive method of ethnically marked and ethnically unmarked neighbourhood was applied. Neighbourhoods with fifty of more per cent of ethnic Poles and/or Russians are referred to as ethnically marked (EthM) in the corpus as well as the present research. Accordingly, Vilnius neighbourhoods inhabited by over sixty per cent of ethnic Lithuanians[6] are called ethnically unmarked neighbourhoods (EthUn) (see Figure 1).
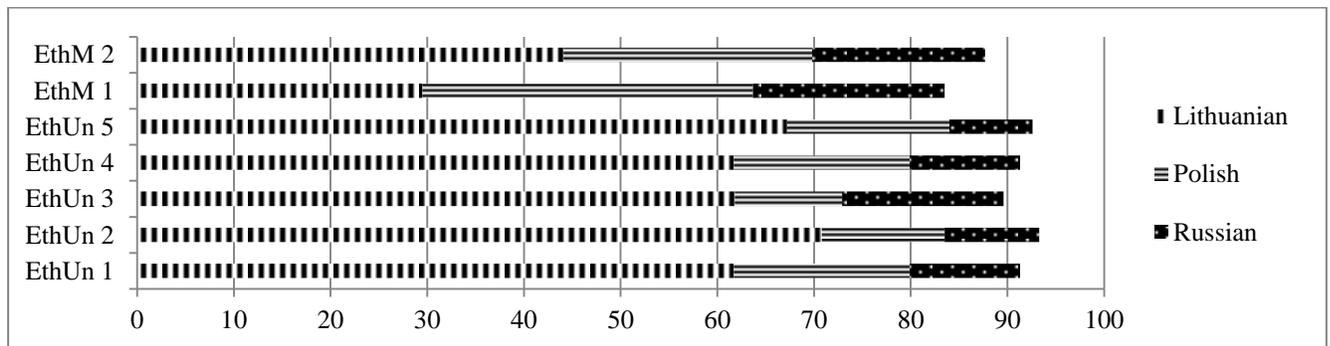

Figure 1. VALC collection: ethnic composition of Vilnius neighbourhoods

---

[6] Results of the 2011 Population and Housing Census of the Republic of Lithuania. Available at: http://statistics.bookdesign.lt/?lang=en. Accessed: 5 September 2014.

As shown in Figure 1, there are more ethnic Poles (approx. 30 %) in EthM 1 and EthM 2 than Russians (approx. 20 %). Ethnicity does not correlate directly with linguistic choices as there is likelihood that adolescents from EthM neighbourhoods may choose Russian in addition to Polish resources. As shown by Geben (2010), Vilnius Poles tend to use both Polish and Russian for public communication on a daily basis. Ethnic Poles sometimes even quarrel in Russian. Hence, Polish or Russian ethnicity does not directly presuppose the use of either Polish or Russian resources.

All the schools in EthUn and EthM neighbourhoods have Lithuanian as the language of instruction. Ethnic minorities in Lithuania have a right to choose schools with minority language of instruction or bilingual schools, e.g. Polish and Lithuanian. Alternatively, adolescents may receive education in schools with Lithuaian as the language of instruction. According to Ramonienė (2013), changes in the socio-political situation of the country after the reestablishment of independence have affected social and linguistic behavior of ethnic minorities. Ethnic Russians tend to send children to Lithuanian rather than Russian schools (ibid.). Ethnic Poles, on the contrary, often send their children to Polish schools (Leončikas 2006), hence, raising the possibility that the schools chosen for the corpus compilation will be attended by very few or even no Poles. It appeared, however, that only slightly less than half of the informants were ethnic Poles, the other half – ethnic Russians (from mixed families, see below).

Obviously, linguistically diverse neighbourhood as well as linguistically diverse family background may correlate with the availability of Russian or Polish resources in adolescents' speech. When filling in the questionnaire, neither of the informants from EthUn neighbourhoods indicated any other language used at home in addition to Lithuanian. Hence, it is likely that adolescents attending shools in EthUn neighbourhoods have Lithuanian as their mother tongue, which may limit their availability to Russian resources. All the informants from EthM neighbourhoods come from ethnically mixed families, namely, with one of the parents Russian or Polish and the other ethnic Lithuanian.

Vilnius, in addition to its linguistic diversity, is also socially diverse, as most urban spaces. An additional criterion of the neighbourhood's social marking was applied[7]. Although, as claimed by Lėgaudaitė (2005), who has compiled the Corpus of Kaunas Teenagers (COKT), social class

---

[7] COLT and COLA are based on the non-linguistic parameters of the social marking of adolescents' residential neighbourhood and the social marking of the school neighbourhood (Stenström, Andersen and Hasund 2002).

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

11

distinctions are much less overt in major cities in Lithuania in comparison to London or any other major European city, some socio-demographic stratification of Vilnius neighbourhoods can be identified (cf. Žilys 2013). For most informants the school neighbourhood is likely to coincide with their residential neighbourhood[8] (Urbanovič 2010) as it is a common and a legitimated practice to choose a school in an adolescent's residential area.

The schools in EthUn neighbourhoods are located in more prestigious neighbourhoods closer to or in Vilnius centre as well as neighbourhoods of the Soviet blocks-of-flats. EthM neighbourhoods, which were mostly inhabited by factory and construction workers during the Soviet period, coincide with working-class districts.

In addition, as schools located in one neighbourhood may be of different status, a contrastive method of a good (high position in the rank of schools) and average school (middle position) was applied[9]. In summary, four schools ranked high and five schools in the middle position were chosen for the project. It appears that the schools ranked in the middle position are located in either the neighbourhoods of the Soviet blocks-of-flats or working class districts, whereas all the schools ranked high are located in or near Vilnius centre.

In summary, VALC has been collected to cover the language of both ethnically diverse Vilnius neighbourhoods as well as ethnically unmarked districts, making it suitable for the present research. Additionally, as Vilnius is far from a socially homogenous city, the social stratification of Vilnius neighbourhoods was attended to, hence, making the corpus representative of a wider range of Vilnius adolescents' language.

3.3 Coding the corpus data: lexico-functional categories of Russian and English resources

English and Russian resources were coded according to the following lexico-functional categories: slang, swearing, code-switching, acronyms, quotations and ready-made phrases. Obviously, these

---

[8] Urbanovič (2010) in the MA paper reveals that approximately 75–80 per cent of school pupils attend schools closest to their place of residence.

[9] The position of the school was identified on the basis of the list of schools ranked according to the results of graduation exams (*Veidas*, 2012-04-23).

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

12

categories were identified on the basis of careful analysis of the transcribed data rather than applied as pre-given category labels.

There is no straightforward, all-covering definition of slang. Definitions range from broad approaches, for example, 'slanguage' (Stenström, Andersen and Hasund 2002: 63) to avoiding the definition altogether by listing what is slang and what is not (cf. Andersson and Trudgill 1990). In this research slang is defined as specific words (with or without morphological adaptation) which are adolescents' in-group related (cf. Lėgaudaitė 2009: 181), e.g. *karočė <EN so, in short>, vsio <EN that's all>, davai <EN let's>, pačekinti <EN check>, laikinti <EN like>, postinti <EN post>, okey, sorry, soriukas <EN dimin. sorry>, feisas <EN face>*[10].

Similarly to slang, there seems to be no all-covering definition of swearing. Scholars approach swearing as a linguistic expression of negative emotions (Wierzbicka 1991). Youth language researchers showed that swearing may index both negative and positive feelings (Fägersten 2012). Swearing in the present paper is perceived in terms of specific words, which are treated as inappropriate, offensive when used outside adolescents' social context but not necessarily offensive in adolescents' peer group, as well as in terms of pragmatic function, i.e. as a solidarity builder (social swearing) or annoyance swearing that occurs in situations of increased stress, as a manifestation or a release of tension (ibid.: 20). For example: *blemba (euphemistic form of the strong swearword blet), idzi nachuj <EN fuck you>, fuck, fucker*.

Although some pieces of research draw on dictionaries as a valuable data for the identification of certain lexical items as slang or swearwords, this approach is not appropriate in the present research. Youth language research suggests that slang or swearing cannot be identified on the basis of their semantics. Rather slang and swearwords gain their value only within the sociocultural context in which they occur. The use and function of linguistic resources in their interactional context is, thus, a fundamental criterion for identifying linguistic resources as slang or swearwords in the present study.

Code-switching is perceived not from a grammatical perspective, but rather as a functional, interactional phenomenon, for example, code-switching is more likely in certain sequential positions or

---

[10] All the data was orthographically transcribed, Russian and English associated resources were transcribed to represent the pronunciation.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

13

is a means for topic management (Auer 2002: 3). Code-switching may occur as longer stretches of talk or individual inserted elements/insertions, e.g. *oi o dieve ką daryt čia tikrai dideli <u>sugrobai</u> <EN oh my god what should we do these are really huge <u>snowpiles</u>>*.

Quotations and ready-made phrases are similar in form – as intertextual references drawing on adolescents' recreational, media practices (cf. Lytra and Baraç 2009). Quotations are produced spontaneously with no external hint, e.g. excerpts from songs, movies, whereas ready-made phrases occur as expressions read from a computer or a smartphone rather than produced spontaneously, e.g. qualities of mobile phones (*no-lag quality, air view*), sentences displayed on a computer screen (*not responding*), computer related actions (*paspausti play <EN press play>*). Finally, acronyms are defined as shortenings which are typical of computer mediated communication, social networking, e.g. *omg* [omg], *lol* [lol], *yolo* [jəuləu].

The lexico-functional categories reveal the functions and significance that English and Russian resources acquire in adolescents' interactional repertoire.

4. Results and discussion

As there may be several possible dimensions of the present analysis, two major aspects have been chosen going from general statistics to the quantitative analysis of specific resources. First, in Section 4.1 I will analyse whether neighbourhood and gender statistically correlate with the frequency of Russian and English resources. Section 4.2 will present the distributional analysis of lexico-functional categories of Russian (4.2.1) and English (4.2.2) resources.

Normalised frequencies of Russian and English resources by gender and ethnic marking of neighbourhoods were calculated with the SPSS statistical program. To identify whether quantitative differences between groups are statistically significant the Mixed ANOVA test was used (results are statistically significant if $p<0.05$).

4.1 Correlation between English and Russian resources and non-linguistic variables

The statistical corpus analysis has shown that ethnic marking of the neighbourhood correlates significantly with the frequency of Russian resources. English resources are distributed differently

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

14

across gender categories in both types of neighbourhoods, whereas Russian resources are distributed differently in boys' and girls' speech in EthUn neighbourhoods only (see Table 2).

Table 2. Distribution of Russian and English resources (tokens and types) across gender groups and neighbourhoods (**= $p<0.01$, *= $p<0.05$, n.s.= no statistical difference)

Resources	Neighbourhoods' ethnic marking	Gender	Average per 1,000 words		p	
			Tokens	Types	Tokens	
Russian	EthM	Boys	61.4	33.5	n.s.	*
		Girls	49.2	12.6		
		Total	53.9	20.6		
	EthUn	Boys	18.7	10.2	**	
		Girls	8.05	4.4		
		Total	12.2	6.7		
English	EthM	Boys	15.3	5.65	**	n.s.
		Girls	3	2.29		
		Total	11.1	3.58		
	EthUn	Boys	10.9	7.0	*	
		Girls	6.1	4.8		
		Total	7.9	5.7		

The tokens of Russian resources predominate over English resources (tokens). In addition, Russian resources used by adolescents from EthM neighbourhoods seem to be more varied in types in comparison to the linguistic repertoire in EthUn neighbourhoods. Hence, availability correlates with more types and tokens of Russian resources.

Different distribution of the tokens of Russian resources has been found across gender categories (in EthUn). Boys display significantly more frequent use of Russian resources in comparison to girls. Following Kiesling's (2009: 177) observation on the relation between statistical co-occurrence and indexical meaning, Russian resources seem to have taken boys' group's indexicality in EthUn neighbourhoods. The correlation between the use of Russian resources and masculinity is less direct in EthM neighbourhoods as there is no statistically significant difference between boys' and girls' use of Russian resources.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

15

Ethnic marking of the neighbourhood is not significant for the frequency of English resources (tokens). Similar distribution of English resources across neighbourhoods should not be generalized as there are exceptions. Girls from EthUn neighbourhoods use twice more English resources (6.1 tokens per 1,000 words) in comparison to girls from EthM neighbourhoods (3 tokens per 1,000 words). English resources for girls in EthUn neighbourhoods may perform the functions that Russian resources play for girls from EthM neighbourhoods (see Section 4.2.1).

Generally, English resources are differently distributed across gender categories in both types of neighbourhoods. Boys use significantly more English resources in comparison to girls. Predominant use of English resources by boys seems contrary to the findings that girls use more English resources than boys (Čekuolytė 2012). It is likely that it is not the overall frequency that distinguishes girls' speech but rather specific categories of resources, e.g. quotations. In addition, as observed from the data collection and analysis, predominant use of English resources by boys relates to computer games, e.g. *geimas <EN game>, spavneris <EN spawner>, šarpas <EN sharp>, skinas <EN skin>*. There is also likely to be correlation between the use of English resources and identity work, e.g. of a gamer (cf. Leppänen (2007) on the use of English expressions for indexing the identity of an experienced gamer).

Russian and English resources are not distributed randomly in Vilnius adolescents' speech but have a systematic organization. Norms of Vilnius adolescents' heteroglossic practices differ across Vilnius neighbourhoods as it appeared that Russian resources are more frequent and more varied in types in EthM neighbourhoods. Gender is also a significant variable for the distribution of both Russian and English resources as both occur more frequently in boys' than girls' speech. The functional significance of Vilnius adolescents' linguistic repertoire is analysed in the following Section.

4.2 Russian and English resources: distribution of lexico-functional categories

The lexico-functional categories of Russian and English resources are distributed differently in EthM and EthUn neighbourhoods as well as in boys' and girls' speech (see Table 3).

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

16

Table 3. Lexico-functional distribution of Russian and English resources (tokens) (%= percentage from the total number of resources, ***= p<0.001, **= p<0.01, *= p<0.05, -= test is unreliable, N= total number of tokens)

| | EthM | | | | | | EthUn | | | | | |
| | Boys | | | Girls | | | Boys | | | Girls | | |
Categories	RU %	p	EN %	RU %	p	EN %	RU %	p	EN %	RU %	p	EN %
Slang	54.0	**	75.0	38.0	***	73.0	60.0	***	80.0	62.0	***	76.0
Code switching	8.0	-	0	35.0	-	0	0	-	0	0	-	0
Swearwords	37.9	***	8.0	25.0	*	16.0	40.0	***	9.9	38.0	***	8.0
Quotations	0.1	***	12.0	2.0	*	6.0	0.0	-	2.0	0.0	-	13.0
Phrases	0.0	-	5.0	0.0	-	5.0	0.0	-	0.1	0.0	-	1.0
Acronyms	0.0	-	0.0	0.0	-	0.0	0.0	-	8.0	0.0	-	2.0
N	704		93	621		102	702		634	509		468

Slang predominates in the language of boys and girls from both types of neighbourhoods. Swearwords rank second. Furthermore, slang correlates with both Russian and English resources, whereas swearing shows stronger correlation with Russian resources. In addition, code-switching relates to Russian resources in EthM neighbourhoods only, showing stronger correlation with girls' speech. Finally, quotations and ready-made phrases correlate with English resources in both types of neighbourhoods. A few instances of acronyms have occurred among adolescents in EthUn neighbourhoods.

In the following Section I will focus on the distributional lexico-functional analysis of Russian resources and exemplify the categories. In Section 4.2.2 I will continue with the lexico-functional analysis of English resources.

4.2.1 Russian resources in VALC: correlation between availability and lexico-functional variation

Russian resources in EthM neighbourhoods show more lexico-functional variation in comparison to EthUn neighbourhoods. There are gender differences, however. Boys' (EthM) language is dominated by Russian slang (*davai <EN come on, let's>, pacanas <EN guy>, zdarova <EN hey>, lievas <EN uncool, boring>*), swearing and several instances of code-switching. In contrast, girls use significantly

fewer Russian swearwords than boys. Girls' speech is dominated by Russian slang and instances of code-switching.

Code-switching appears significant both due to its frequency in comparison to other lexico-functional categories as well as due to its absence among adolescents in EthUn neighbourhoods. The degree to which Russian resources are drawn by girls and boys from EthM neighbourhoods varies from occasional insertions (example (1)) to longer utterances (2), for example:

(1)   (Girl, EthM)
aš su juo tik <u>pamirylas</u> taip ir parašysiu nes nežinau[11]
<EN I have just <u>reconciled</u> with him so I am going to write so>

 (2) (Boys and a girl, EthM)

1	B1:	<u>*nu vo kak v škole*</u>
		<EN <u>how is it at school?</u>>
2	G1:	*nu <u>chorošo kak u tebe</u>*
		<EN well <u>ok and what about you?</u>>
3	B1:	<u>*ai vsivo byvajit*</u>
		<EN well <u>everything happens</u>>
4	G2:	<u>*kak ocenočki*</u>
		<EN <u>how are the grades</u>>
5	B1:	<u>*oi daže lutše nesprašivai*</u> *karočia pas mus tokia nesveika mokytoja nu taip aš*
6		*jos nekenčiu*
		<<u>you'd better not ask</u> you know we have such a dumb teacher I hate her so much>
7	G2:	*na būna ir tokių bet reikia*
		<EN well there are such teachers but they are necessary>
8	G1:	*nebuvo fizikos tai padarė dvi matematikas ačiū dievui matematika turėjo būt ketvirta ir tipo*
9		*šešta paskutinė*
		<EN we didn't have physics so they made two mathematics thanks God mathematics had to be the fourth and you know the sixth the last>

The interactional sequence in extract (2) appeared at the beginning of the recording. Code-switching into Russian in lines 1–5 serve as resources for polite socializing. In line 5 the topic of school achievement is touched upon, which provokes irritation and dissatisfaction (marked by faster tempo of speech). This switch to Lithuanian appears interactionally significant, marking a shift in topic and emotional attitude of speaker.

Example (3) further illustrates the significance of Russian resources as instances of code-switching:

---

[11] Russian resources are underlined.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

18

(3) (Boys and a girl, EthM)

1   B2:   *eee netriukšmauk.*
               <EN eee do not make noise>
2   M1:   *xxx[12] ne*
               <EN xxx no>
               %sit: $[13]
3   B2:   *eee š*
               <eee shhh>
4   B2:   *eik išsaugok*
               <EN go save it>
5   B1:   *tai būtų <u>u tsibia xxx zadnica kvadratnaja</u>*
               <EN it would be <u>your xxx ass is square</u>>
               %sit: $
6   B2:   *ne vis tiek jis traumavo*
               <EN no anyway he harmed>

The local interactional context in example (3) focuses on playing a computer game. It is, however, not clear which game is being played as the conversational excerpt seems more directed to negotiating roles and arguing rather than the game. B1 in line 5 switches to Russian for a provocative insult which is intended as opposing the order not to make too much noise in line 1. It appears, however, that the speaker's attempt to resist the order and sound witty is not responded to (B2 might not have heard the insult as he was speaking simultaneously with the utterance in line 5). It is likely that B1's linguistic choice in line 5 is strategic. The insult in Russian appears to be related to an attempt to negotiate social status and dominance, in other words, the boy who can give orders and insult has hierarchy within the group.

Code-switching, although illustrated by several examples only, is employed in meaningful ways by participating speakers. Consequently, Russian resources can be perceived by speakers as distinct codes (see Section 2 on the notion of 'code'). Namely, they may provide possibilities for topic management as well as serve as markers of we-code (for socializing, establishing and negotiating peer group status). In other words, code-switching is a resource for adolescents in ethnically marked neighbourhoods.

Russian quotations among girls from EthM neighbourhoods serve as a stylistic game, a means of a playful expression. For example:

---

[12] *xxx* – unintelligible speech
[13] *%sit: $* – overlapping speech

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

19

(4) (Girl, EthM)
<…> nichačiu učytsa, chačiu ženytsa
<EN I don't want to study, I want to get married>

(5) (Girl and boy, EthM, when playing cards)
1   B28:   *pas mane dvoika.*
        <EN I have two>
2   G29:   *o tą parodyk.*
        <EN show that one>
3   G26:   *parodyk savo dvoiką da ladna prapuskaim poniat i prostit žinai kaip mes*
        <EN show me your two ok let it be you may pass to understand and to forgive you know as we>

 (6) (Boy, EthM)
<…> jou mama rodnaja što mne chuligana na na na na nimerena bet ką čia rusiškai mes kalbam
<EN jou my momwhat for me a hooligan na na na na but why are we speaking here in Russian>

The quotations in (4) and (5) have occurred as insertions in conversations. Availability to Russian resources by girls (EthM) might as well correlate with local interactional purposes to sound creative and playful. In (4) the quotation is used to convey a relaxed attitude to life. *Poniat i prostit* in (5) serves as a comment during a game of cards. An excerpt from a Russian rap song in (6) indexes the importance of Russian pop-culture for adolescents in EthM neighbourhoods. It seems that adolecents (EthM) position themselves as members of the Russian culture. In addition, the metalinguistic comment on using Russian in the conversation in (6) shows that adolescents (at least in EthM neighbourhoods) identify resources as Russian. In general, adolescents from EthM neighbourhoods have Russian resources available and they design their own linguistic uses, which they find meaningful and functional.

Not surprisingly, no instances of Russian resources as code-switching have been observed in the data from EthUn neighbourhoods (see Table 3). Russian resources among adolescents in EthUn neighbourhoods seem to be limited to slang and swearing, with slang predominating. As the categories are similarly distributed in the corpus data of boys' (60% slang and 40% swearwords) and girls' (62% slang and 40% swearwords) speech, it is twice more frequent use of Russian slang and swearwords that indexes masculinity among boys from EthUn neighbourhoods, i.e. 18.7 tokens per 1,000 words by boys and 8.05 tokens by girls. Stenström and Jørgensen (2009: 77) claim that the stronger the group affinity, the more swearing. It could be deduced that boys (in both EthM and EthUn) mark their adherence to

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

20

boys' peer group by frequent use of Russian swearwords. Hence, Russian slang and swearing seem to correlate with the sociolinguistic variable of gender[14].

To exemplify the most common Russian resources in VALC (EthM and EthUn), fifteen most frequent instances, which include either slang (sl) or swearwords (sw), have been identified (see Table 4).

Table 4. Most frequent Russian resources by gender and ethnic marking of the neighbourhood (sw= swearword, sl= slang, %= percentage from the total number of Russian resources)

No	Ethnic marking of the neighbourhood		EthM		EthUn	
Most frequent Russian resources		Total no	Boys % (total number)	Girls % (total number)	Boys % (total number)	Girls % (total number)
1	*blet (sw)*	299	**0.5 (4)**	**5.3 (31)**	**17 (154)**	**19.8 (110)**
2	*karočia (sl) <EN so, in short>*	248	**0.5 (4)**	**10.7 (62)**	**11.1 (100)**	**14.7 (82)**
3	*blemba (sw) (euphemistic form of blet)*	216	**2.5 (18)**	**8.3 (48)**	**4.6 (42)**	**19.4 (108)**
4	*davai (sl) <EN let's, come on>*	157	**1.4 (10)**	**6.7 (39)**	**4.9 (45)**	**11.3 (63)**
5	*čiūju (sl) <EN (most) probably>*	68	0.005 (3)	-	0.072 (33)	0.057 (32)
6	*vapščė (sl) <EN at all>*	50	**0.4 (3)**	**3.1 (18)**	**1.2 (11)**	**3.2 (18)**
7	*zajabys (sl) <EN cool>*	37	-	-	0.072 (33)	0.007 (4)
8	*blyn (sw) (euphemistic form of blet)*	18	0.009 (5)	0.015 (7)	-	0.01 (6)
9	*prikolas (sl) <EN fun, unexpected, funny situation>*	18	-	-	0.019 (9)	0.016 (9)
10	*čystai (sl) <EN absolutely, really>*	17	-	-	0.019 (9)	0.014 (8)
11	*čiuvakas (sl) <EN guy>*	17	0.003 (2)	-	0.01 (5)	0.018 (10)
12	*pofik (sw)<EN no matter, I don't care>*	16	-	0.036 (16)	-	-
13	*lievas(-ai)(sl) <EN uncool, boring, uninteresting>*	11	0.009 (5)	-	0.013 (6)	-
14	*prikin (sl) <EN imagine>*	11	-	0.011 (5)	-	0.01 (6)
15	*paliubomu (sl) <EN really>*	3	0.005 (3)	-	-	-

There are more similarities between boys' and girls' slang and swearwords used in EthUn neighbourhoods (10 slang items and swearwords overlap), whereas there are fewer instances of Russian resources (6 overlap) common for both boys and girls from EthM neighbourhoods. Similarities in the linguistic repertoire of boys and girls from EthUn neighourboods are not surprising as adolescents

---

[14] Similar observations on the correlation of more frequent use of slang, swearing and masculinity made by Stenström (2003: 106), Lėgaudaitė (2009: 188).

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

21

move through different EthUn neighbourhoods, in and out school environments and, hence, the norms for using Russian resources are similar. Due to supposedly higher access to the Russian resources (for many of the students Russian might be a home language) in EthM neighbourhoods the adolescents use them to create various stances and for individual purposes in an interaction and, hence, Russian resources need not overlap in boys' and girls' language.

Girls (EtnM and EthUn) tend to use the strong swearword *blet* more frequently than boys. This observation cannot be generalized, but it is in line with Stenström, who states that taboo expressions that used to belong to "the restricted speech of male subculture <…> are today losing their exclusiveness and isolation" (2003: 74). In addition, it is likely that girls cross (Rampton 2010) into boys' linguistic practices for temporarily negotiating in-group member's identity. Hence, the evaluative cliché of frequent and dirty boys' and tidy girls' language should be abandoned (cf. Forsskähl 2002; Fägersten 2012).

In summary, Vilnius adolescents have a repertoire of Russian resources, which are used differently in EthM and EthUn neighbourhoods as well as across gender categories. Availability to Russian resources in EthM neighbourhoods correlates with wider linguistic repertoire and, hence, more functions. This association of Russian resources in EthM neighbourhoods with lexico-functional variation, different functions in boys' and girls' speech amounts to the existence of norms for adolescents' linguistic practices in EthM neighbourhoods. These norms are different among adolescents in EthUn neighbourhoods, where Russian resources are limited to slang and swearing with fewer gender differences than EthM neighbourhoods.

4.2.2 English resources in VALC: lexico-functional variation

Adolescents encounter English resources on a daily basis and in various modes of communication, i.e. from computer mediated communication and texts of popular culture. This multimodality, or rather transmodality, according to Blommaert (2010), of English resources is reflected in greater lexico-functional variation in comparison to Russian resources.

It has been mentioned that the majority of English resources have been categorized as slang, e.g. *pačekinti <EN check>, laikinti <EN like>, postinti <EN post>, okey, sorry, soriukas<EN sorry,*

diminutive>. Slang is similarly distributed among adolescents in both types of neighbourhoods (see Table 3).

English quotations occurred in the language of adolescents in both EthM and EthUn neighbourhoods as excerpts from songs, films and clichéd quotations, for example:

(7)   (G, EthUn)
        i want to be such a big star i want to shine like a sun want success i want stardome a life fo- a life of fortune and fame

(8)   (G, EthM)
        I'm a lover I'm a mother I'm a sinner

(9)   (B, EthUn)
        titi tiririri in every day you wanna be alright tiririri

(10)   (B, EthM)
        *ir taip toliau taip padaro rankas ir taip po to ant pačios pabaigos **i dont fucking know someone why do you keep asking this questions**[15]*
        *<EN and then he puts his hands this way and then at the very end **i dont fucking know someone why do you keep asking this questions**>*

Quotations from songs were used by both boys and girls as insertions, indirectly related to the content of the interaction, except example (10), which occurred when narrating a film plot.

The correlation between English quotations and gender is not straightforward (see Table 3). The use of English quotations correlates girls' (EthM) language, hence, supporting previous research by Čekuolytė (2012). In contrast, boys from EthM neighbourhoods use twice more English quotations than girls (EthM). The latter observation cannot be generalized as the frequency of English quotations has been influenced by one conversation, namely, frequent use of quotations from one film by two boys.

In addition, adolescents reproduced ready-made phrases, which were read from the Internet or a smart phone rather than produced spontaneously, for example:

(11)   (B, EthUn)
        *va šitas va kur bus tai paspausti **play***
        *<EN this one where this will be so press **play**>*

---

[15] English resources are in bold

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

23

(12) (B, EthUn)
  *keista nu pažiūrim youtube'ą google chrome'as **not responding***
  *<EN strange we look at youtoube google chrome **not responding**>*

(13) (B, EthM)
  *geras sms'as piau jūs gavote sms'ą piau arba emmm **you have a new sms message***
  *<EN cool sms piau you got an sms piau eeem **you have a new sms message**>*

There seem to be no quantitative differences in the distribution of English ready-made phrases in the language of adolescents from both types of neighbourhoods (see Table 3), which hints towards the correlation between English resources appropriated from the mass media, digital technology and adolescents' language, irrespective of the neighbourhood's ethnic marking.

Finally, several acronyms have been found in the corpus data of adolescents from EthUn neighbourhoods, e.g. *yolo* pronounced as [jəuləu] *(you only live once), omg* [omg] *(oh my god)* and *lol* [lol]. The use of *lol* is ambiguous since it may carry the meaning 'laugh out loud' or 'lots of love'. The ambiguity cannot be resolved by sequential analysis as the speaker's use of the acronym does not receive any reaction from the interlocutors. This may suggest that the speaker has used the acronym as language play, as fun. In addition, it appears that adolescents may mix resources from different languages as well as different domains of use, i.e. spontaneous speech in-group slang, computer mediated communication or instant messaging. It is likely that such examples illustrate a playful, creative, hybrid aspect of adolescents' linguistic practices.

When the most frequent Russian and English resources are compared, it becomes apparent that adolescents (EthM and EthUn) have more Russian resources in common than English (see Table 4 and Table 5).

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

24

Table 5. Most frequent English resources by gender and ethnic marking of the neighbourhood (sw= swearword, sl= slang, %= percentage from the total number of Russian resources)

No	Ethnic marking of the neighbourhood		EthM		EthUn	
	Most frequent English resources	Total no	Boys % (total number)	Girls % (total number)	Boys % (total number)	Girls % (total number)
1	*okay (sl)*	94	-	0.52 (23)	0.08 (40)	0.05 (31)
2	*fuck (sw)*	60	**5 (6)**	**18.7 (18)**	**2.2 (12)**	**4 (24)**
3	*laikinti <EN like> (sl)*	36	**1.6 (2)**	**17.7 (17)**	**0.7 (4)**	**2.1 (13)**
4	*čytas/čytinti <EN cheat> (sl)*	22	-	-	0.03 (18)	0.007 (4)
5	*sorry (sl)*	22	-	-	-	0.03 (22)
6	*yes (sl)*	22	-	-	0.03 (17)	0.009 (5)
7	*džokeris <EN joker> (sl)*	13	0.28 (13)	-	-	-
8	*what (sl)*	13	-	-	0.01 (5)	0.009 (5)
9	*geimas <EN game> (sl)*	11	-	-	0.02 (11)	-
10	*nice (sl)*	10	-	-	-	0.01 (10)
11	*cool (sl)*	8	-	-	-	0.01 (8)
12	*flešiukas <EN flash> (sl)*	6	-	-	-	0.01 (6)

Two instances of English resources (*fuck* and *laikinti*) overlap in the language of boys and girls from both types of neighbourhoods. There are differences in frequency, however. Girls use *fuck,* which is weaker than Russian swearwords, twice more (EthUn) or even three times more (EthM) often than boys, which suggests that the English swearword does not relate to masculinity (Čekuolytė 2014). Girls use more instances of *laikinti <EN like>* than boys, which can be explained by their preoccupation with talking about liking photos on Facebook or other social networks.

Boys from EthUn neighbourhoods use computer games related slang (*čytas <EN cheat>* and *geimas <EN game>*) as well as *yes*, the latter used for indexing positive evaluation of the move during the game. In addition, *what* was used as a reproach to the computer for the player's unsuccessful move. *Sorry, nice* and *cool* were used as insertions, evaluative markers among girls from EthUn neighbourhoods.

There seem to be correlations between English resources and computer related activities, social networking on the Internet as well as indexing positive stance towards the listener. Furthermore, it is likely that the speaking situation rather than gender correlates with the choice of English resources.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

25

5. Conclusions

Vilnius adolescents' use of Russian and English resources is a norms-governed linguistic practice. These norms vary along the parameters of gender and ethnic marking of Vilnius neighbourhoods, as evidenced by different frequency of Russian and English resources, different distribution of lexico-functional categories.

Availability to Russian resources correlates with both frequency and variation of lexico-functional categories. Russian resources among adolescents from EthM neighbourhoods are lexico-functionally manifold, i.e. slang, swearing, code-switching and quotations. Russian resources may be used for negotiating hierarchical status in a peer group (namely, code-switching) as well as displaying various stances in interaction (code-switching and quotations), for example, playfulness, creativity, affiliation with the Russian pop-culture. In other words, adolescents in EthM neighbourhoods develop and negotiate their own norms for using Russian resources that they have available on a daily basis.

There is less variation in the use of Russian resources in EthUn neighbourhoods. Russian resources among adolescents (EthUn) are limited to slang and swearing. Although the majority of Russian slang items and swearwords in EthUn neighbourhoods are the same in boys' and girls' speech, there is a norm for boys to use Russian resources twice more frequently as a marker of toughness and masculinity.

Not surprisingly, English resources, which are equally available to adolescents in both types of neighbourhoods, are more varied lexico-functionally in comparison to Russian resources. English resources in EthM and EthUn neighbourhoods, in boys' as well as girls' speech are associated with adherence to pop-culture, computer mediated communication, technological advancements. The available English resources are locally employed and and combined to serve interactional purposes, for mutual entertainment and fun. English resources appeared non-repetitive, varied in types, which suggests that the norms of using them may be peer-group specific or correlate with specific situational contexts rather than the broader category of gender as, for example, frequent use of Russian slang and swearwords correlates with by boys' (EthUn) speech.

The conclusions drawn from the present research contribute to the central points of discussion raised by youth language researchers on adolescents' linguistic practices in linguistically, socially, ethnically heterogeneous urban societies. The functional variety of linguistic resources in Vilnius adolescents' speech confirms the view of strategic, purposeful, creative aspects of adolescents' heteroglossic practices.

Acknowledgements

References

Androutsopoulos, J. K. and A. Georgapoulou (eds.). 2003. *Discourse Constructions of Youth Identities.* Amsterdam: John Benjamins Publishing Company.

Andersson, L. G. and P. Trudgill. 1990. *Bad Language.* Oxford: Blackwell.

Auer, P. 2002. *Code-switching in conversation. Language, interaction and identity.* London: Routledge.

Blommaert, J. 2005. *Discourse. A Critical Introduction.* Cambridge: CUP.

Blommaert, J. 2010. *The Sociolinguistics of Globalization.* Cambridge: CUP.

Blommaert, J. and B. Rampton. 2011. Language and Superdiversity. *Language and Superdiversities* 13 (2), 1–22.

Blommaert, J. 2013. Citizenship, Language and Superdiversity: Towards Complexity. *Journal of Language, Identity and Education* 12, 193–196.

Bucholtz, M. 2009. From Stance to Style: Gender, Interaction, and Indexicality in Mexican Immigrant Youth Slang. *Stance: Sociolinguistic Perspective.* Jaffe, A. (ed.). Oxford: OUP. 146–170.

Cáccamo, C. A. 2002. From 'switching code' to 'code-switching': Towards a reconceptualisation of communicative codes. *Code-Switching in Conversation. Language, Interaction and Identity.* Auer, P. (ed.). London: Palgrave. 29–50.

Čekuolytė, A. 2012. Etnografinio metodo taikymas sociolingvistiniuose (jaunimo kalbos) tyrimuose. *Taikomoji kalbotyra* 1, 1–16. http://taikomojikalbotyra.lt/lt/2012/04/etnografinio-metodo-taikymas-sociolingvistiniuose-jaunimo-kalbos-tyrimuose.

Čekuolytė, A. 2014. *He blet nachui was in a shop*: Swearing Practices and Attitudes to Swearing among Vilnius Adolescents. *Taikomoji kalbotyra* 6. http://taikomojikalbotyra.lt/lt/numeris/2014-6.

Forsskähl, M. 2001. Girls' slang and boys' slang: two towns in Finland visited. *Nordic Journal of Women's Studies* 2 (9), 98–106.

Fägersten, K. B. 2012. *Who's Swearing Now? The Social Aspects of Conversational Swearing.* Cambridge: Cambridge Scholars Publishing.

Geben, K. 2010. Vilniaus lenkų kalbinė savimonė, daugiakalbystė ir tapatybė. *Miestai ir kalbos.* Ramonienė, M. (ed.).Vilnius: Vilniaus universiteto leidykla. 153–173.

Jørgensen, J. N., Karrebæk, M. S., Madsen, L. M. and J. S. Møller. 2011. Polylanguaging in Superdiversity. *Language and Superdiversities* 13 (2), 23–39.

Kiesling, S. F. 2009. Style as Stance: Stance as the Explanation for Patterns of Sociolinguistic Variation. *Stance. Sociolinguistic Perspectives.* Jaffe, A. (ed.). Oxford: OUP. 171–194.

Leončikas, T. 2006. Asimiliacijos prielaidos ir mokyklos pasirinkimas: pirmokų tėvų grupių palyginimas. *Sociologija. Mintis ir veiksmas* 1, 76–98.

Leppänen, S. 2007. Youth language in media contexts: insights into the functions of English in Finland. *World Englishes* 26 (2), 149–169.

Lėgaudaitė, J. 2005. Age-Specific and Gender-Specific Slang. *Baltu Filoloģija XIV* 2, 77–88.

Lėgaudaitė, J. 2009. Similarities and Differences between slang in Kaunas and London Teenagers' speech. *Youngspeak in a Multilingual Perspective.* Stenström, A. B. and A. M. Jørgensen (eds.). Amsterdam: John Benjamins Publishing Company. 177–203.

Lytra, V. and T. Baraç. 2009. Multilingual practices and identity negotiations among Turkish speaking young people in a diasporic context. *Youngspeak in a multilingual perspective*. Stenström, A. B. and A. M. Jørgensen (eds.). Amsterdam/Philadelphia: Benjamins. 55–78.

Makoni, S. and A. Pennycook (eds.). 2007. *Disinventing and Reconstructing Languages.* Clevedon: Multilingual Matters.

Nevinskaitė, L. 2010. Kalbų mokėjimas ir vartojimas didžiuosiuose miestuose: lyginamoji analizė. *Miestai ir kalbos.* Ramonienė, M. (ed.). Vilnius: Vilniaus universiteto leidykla. 269–277.

Otsuji, E. and A. Pennycook. 2009. Metrolingualism: fixity, fluidity and language in flux. *International Journal of Multilingualism* 7 (3), 240–254.

Ramonienė, M. 2013. Family Language Policy and Management in a Changed Socio-political Situation: Russians and Russian Speakers in Lithuania. *Successful Family Language Policy. Multilingual Education 7.* Schwartz, M. and A. Verschik (eds.). Dordrecht: Springer. 127–143.

Rampton. B. 2010. From 'multi-ethnic urban heteroglossia' to 'contemporary urban vernaculars'. *Urban Language and Literacies* 61, 2–24.

Stenström, A. B., Gisle, A. and I. K. Hasund. 2002. *Trends in Teenage Talk. Corpus compilation, analysis and findings.* Amsterdam: John Benjamins Publishing Company.

Stenström, A. B. 2003. *It's not that I really care, about him personally you know:* The construction of gender identity in London teenage talk. *Discourse Constructions of Youth Identities.* Andoroutsopoulos, J. K. and A. Geograpoulou (eds.). Amsterdam: John Benjamins Publishing Company. 93–119.

Stenström, A. B. and A. M. Jørgensen. 2009. *Youngspeak in a Multilingual Perspective.* Amsterdam: John Benjamins Publishing Company.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

Urbanovič, R. 2010. *Vilniaus miesto ir rajono vidurinių mokyklų ir mokinių kaita 1990 – 2008 m.* Lithuanian University of Educational Sciences. Unpublished MA thesis.

Vaicekauskienė, L. 2010. Globalioji daugiakalbystės perspektyva: anglų kalbos vieta ir vaidmuo Lietuvos miestų erdvėje. *Miestai ir kalbos.* Ramonienė, M. (ed.). Vilnius: Vilniaus universiteto leidykla. 175–203.

Vaicekauskienė, L. 2014. Transformations of social values of language and stagnation of language policy in post-soviet Lithuania. *Europe Regional* 3 (4). [In print]

Verschik, A. 2010. Contacts of Russian in the post-Soviet space. *Applied Linguistics Review* 1, 85–128.

Verschueren, J. 1991. A Pragmatic perspective on international communication. *The Pragmatics of International and Intercultural Communication.* Blommaert, J. and J. Verschueren (eds.). Amsterdam: John Benjamins Publishing. 187–210.

Vyšniauskienė, I. 2012. Polylingual Behaviour of Vilnius Youth: Searching for social identity. *Kalbotyra* 64, 140–157.

Wierzbicka, A. 1997. *Understanding Cultures Through Their Key Words: English, Russian, Polish, German and Japanese.* New York: Oxford University Press.

Žilys, A. 2013. Rezidencinė diferenciacija ir skirtumai Lietuvos moderniajame mieste: (po)sovietinis ar Vakarų miestas? *Kultūra ir visuomenė* 4 (2), 67–101.

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

30

Rusiški ir angliški ištekliai Vilniaus paauglių kalboje: tekstynu paremtas tyrimas

Inga Vyšniauskienė

Santrauka

Didmiesčiuose daugėjant daugiakalbių gyventojų, intensyviau pradėta tyrinėti ir didmiesčiuose vartojama jaunimo kalba. Tyrėjai pastebi, kad šiuolaikinis jaunimas, norėdamas savo kalbinei veiklai suteikti kūrybiškumo ir žaismingumo, dažnai savo kalboje įterpia ir kitų kalbų leksikos, pvz., danai paaugliai vartoja turkų, britai – vokiečių kalbos išteklių.

Teorinį jaunimo kalbos tyrimų pagrindą sudaro postruktūralistinis požiūris į kalbą kaip sociokultūrinį konstruktą. Manoma, kad kalbos kaip uždaros sistemos samprata nėra tinkama šiuolaikinei įvairialypei miesto paauglių kalbinei vartosenai apibūdinti: paaugliai ne tik įterpia kitų kalbų leksikos, bet drauge maišo ir fonetinius, sintaksinius įvairių kalbų bruožus, todėl atskirų kalbų ribų dažnai neįmanoma nubrėžti. Jaunimo kalbai apibūdinti tyrėjai pasitelkia kitai kalbai būdingų išteklių sąvoką.

Darant prielaidą, kad Vilnius kaip daugiakalbis Lietuvos miestas yra itin tinkama terpė rastis įvairiakalbei, mišriai vartosenai, tyrimui pasirinkta būtent sostinės jaunimo spontaniška kalba. Tyrimo empirinis pagrindas - beveik 170 tūkst. žodžių Vilniaus paauglių (10–16 m.) iš skirtingos socialinės ir etninės sudėties Vilniaus rajonų kalbos tekstynas. Etniškai žymėtais rajonais įvardyti tie, kuriuose didžioji gyventojų dalis yra lenkai ir rusai, etniškai nežymėtais – kuriuose daugiausia lietuvių tautybės gyventojų. Visose tekstyne pasirinktose mokyklose dėstomoji kalba yra lietuvių.

Kiekybinės analizės duomenimis, paaugliai iš etniškai žymėtų rajonų vartoja statistiškai daugiau ir įvairesnių rusų kalbos išteklių. Be to, jiems būdinga didesnė leksinių ir funkcinių rusų kalbos išteklių kategorijų įvairovė – ne tik rusų slengas ir keiksmažodžiai (kaip ir paaugliams iš etniškai nežymėtų rajonų), bet ir citatos, kodų kaita. Etniškai nežymėtuose rajonuose vaikinai vartoja du kartus daugiau rusų kalbos slengo ir keiksmažodžių išteklių nei merginos, todėl tikėtina, kad rusų kalbos ištekliai koreliuoja su vyriškumu.

Anglų kalbos išteklių vartosenai etninis rajono žymėtumas įtakos neturi. Gali būti, kad tą lemia skirtinga prieiga prie anglų kalbos resursų: tiek žymėtų, tiek nežymėtų rajonų paaugliai su anglų kalbos

Vyšniauskienė, I. Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech, *Taikomoji kalbotyra* 2014 (6), www.taikomojikalbotyra.lt

31

ištekliais susiduria per populiariąją kultūrą, internetą. Taip pat tyrimas parodė, kad Vilniaus paauglių kalba nėra chaotiškas kitakalbių išteklių kratinys, tačiau savitas normas turinti kalbinė praktika.

Corpus Analysis of Russian and English Resources in Vilnius Adolescents' Speech

Inga Vyšniauskienė

Summary

This paper presents a corpus analysis of Vilnius adolescents' heteroglossic practices, namely the use of Russian and English resources. Vilnius (the capital of Lithuania) has been chosen as a fruitful niche for the present research as it is Lithuania's most linguistically and culturally diverse city.

The approach to adolescents' linguistically diverse speech as a meaningful practice conceptually contrasts the viewpoint commonly voiced in society at large on the inadequate mixture, unnecessary insertion of Russian and English words. The present analysis is intended as a scientific argument for the existence of norms for using Russian and English resources among Vilnius adolescents. Namely, these norms will be investigated along the parameters of gender and ethnic marking of Vilnius neighbourhoods.

The analysis draws on Vilnius Adolescents' Language Corpus data, collected for the present research, documenting the speech of 10–16 year old Vilnius adolescents in ethnically marked (EthM), i.e. predominantly populated by ethnic Poles and/or Russians, and ethnically unmarked (EthUn), i.e. predominantly populated by ethnic Lithuanians, Vilnius neighbourhoods. All the schools have Lithuanian as the language of instruction.

The research has shown that Vilnius adolescents' linguistically diverse practices are far from a mere mixture of resources. Instead, adolescents have different norms for using Russian and English resources.

Adolescents in EthM neighbourhoods appear to use more types as well as tokens of Russian resources, in comparison to adolescents from EthUn neighbourhoods, hence, availability correlates with usage. The lexico-functional categories of Russian resources among adolescents (EthM) are manifold, i.e. slang, swearing, code-switching and quotations. Russian resources in the language of adolescents from

EthUn neighbourhoods are limited to slang and swearing. Boys (EthUn) use Russian resources twice more frequently than girls, hence, it is likely that Russian slang and swearwords serve as markers of toughness and masculinity.

English resources are available to adolescents irrespective of the ethnic marking of the neighbourhood through computer-mediated communication, popular culture. Hence, their frequency and lexico-functional categories are similarly distributed in Vilnius adolescents' linguistic practices. It appeared that English resources are often situation, topic specific, employed for mutual entertainment and fun.

Įteikta 2014 gruodžio mėn.