Michał Głuszkowski
Nicolaus Copernicus University in Toruń
(Univerza Nikolaja Kopernika v Torunju)
micglu@umk.pl

Code-Switching and Code-Mixing on Different Language Levels: The Case of a Polish Language Island in Siberia

The aim of the article is to discuss the diverse nature of code-switching and mixing phenomena on different language levels. In the field of phonetics, morphology and lexis, researchers most often focus on the interlingual influence in the form of interference, while code-switching is a phenomenon that occurs at the level of syntax. In this article about linguistic material from a Polish language island in a Russian environment (the village of Vershina in Siberia), examples of lexical and structural borrowings and various forms of language change during speech were indicated. However, this approach only allows for the characterization of selected fragments of utterances, while whole texts function in live communication. Therefore, it was proposed to include the most complex level of language, i.e., the text as such, in the analyses of code-switching and mixing. Selected examples show how this complementation affects the perception of phenomena related to bilingualism.[1]

Keywords: code-switching, code-mixing, language contact, bilingualism, sociolinguistics

Preklapljanje in mešanje kodov na različnih jezikovnih ravneh: primer poljskega jezikovnega otoka v Sibiriji

Cilj prispevka je obravnavati raznolikost pojavov preklapljanja in mešanja kodov na različnih jezikovnih ravneh. Na področju fonetike, morfologije in leksike se raziskovalci najpogosteje osredotočajo na medjezikovne vplive v obliki interference, medtem ko je kodno preklapljanje pojav, ki se pojavlja na ravni skladnje. V članku o jezikovnem gradivu s poljskega jezikovnega otoka v ruskem okolju (vas Veršina v Sibiriji) so bili nakazani primeri leksikalnih in strukturnih izposojenk ter različne oblike jezikovnih sprememb med govorom. Vendar ta pristop omogoča le karakterizacijo izbranih fragmentov izrekov, medtem ko v živem sporazumevanju delujejo celotna besedila. Zato je bilo predlagano, da se v analize kodnega preklapljanja in mešanja vključi najkompleksnejša jezikovna raven, tj. besedilo, izbrani primeri pa pokažejo, kako to dopolnjevanje vpliva na dojemanje pojavov, povezanih z dvojezičnostjo.

Ključne besede: kodno preklapljanje, kodno mešanje, jezikovni stik, dvojezičnost, sociolingvistika

1 Introduction

1.1 The community under study

The issue of code-switching and mixing is one of the most discussed topics in the literature on language contacts. The purpose of this article is to analyse the phenomenon in question on different language levels in a community that emerged as a result of the

---

migration of Slavic community and their settlement in a foreign, but also Slavic-speaking environment. Vershina is a village founded by Polish migrants in Russia in 1910 in the Irkutsk oblast. Unlike most Poles in Siberia, they were not exiles, but voluntary settlers from various parts of Lesser Poland, who wanted to improve their material situation by taking advantage of the land grants and tax reliefs made possible by the land reform introduced by the Prime Minister of the Russian Empire Peter Stolypin in 1906 (Bazylow 1975: 72). The community was bilingual from the beginning, because the settlers came from the areas that were part of the Russian partition (Polish lands incorporated into Russia in the years 1795-1918). Their bilingualism is connected with diglossia, which was initially characterized by the dominance of the Lesser Poland dialect in most domains, and currently it is used to a greater extent only in ethnically homogeneous Polish families (Głuszkowski 2012: 46-7).

Vershinian community represents the "folk" and asymmetrical type of bilingualism, which means that one of the languages has been acquired in the course of daily communication, and the knowledge and use of both codes is unbalanced – with a clear dominance of the majority language, i.e. the official language of the country (cf. Bullock, Toribio 2009: 9; Głuszkowski 2015: 56).

The material under analysis was collected in the course of field research during 4 expeditions to Vershina and includes almost 100 hours of individual and group interviews. A significant part of the material was recorded and allowed for the creation of a corpus of 275,000 words.

1.2 The specificity of language contact in language islands

The community in which the linguistic material analysed in this study was obtained can be characterized as an insular settlement. According to Claus Hutterer, "language islands are internally structured settlements of a linguistic minority on a limited geographical area in the midst of a linguistically different majority" (cit. and transl. after Rosenberg 2005: 221). Thus, the island must be clearly smaller than the sea on which it is located (Nowicka 2011), and it is a specific kind of minority. Although this term (Ger. 'Sprachinsel') was coined to characterize a Slavic community surrounded by a German majority in East Prussia over 170 years ago, it is still widely used in various subdisciplines of linguistics (Rosenberg 2005: 221). One of the main reasons to study language islands was their isolation, because of which the researchers expected to analyse phenomena of language contact as well as preserved archaisms in an undisturbed form (cf. Löffler 1987: 387). However, for the same reason the concept of island has been criticised, e.g. by Erik Eriksen, who argues that neither a community is entirely isolated, nor cultural boundaries are absolute, because "webs of communication and exchange tie societies together everywhere, no matter how isolated they may seem at a first glance" (Eriksen 1993: 134). This does not mean, however, that this idea should be rejected, but treated with appropriate reserve, as an analytical model based on relative, not absolute isolation. It was in this spirit that the concept of island communities was

developed in sociolinguistics, especially in relation to languages with a small number of speakers:

> Островные ситуации возникают вследствие переселения части того или иного этноса по тем или иным причинам – социально-политическим, военным, экономическим и прочим – в регионы проживания иного этноса (иных этносов) с иным языком (иными языками). Оторванные от исходного этно-языкового корня, такие острова постоянно ощущают языковой дефицит, особенно что касается его использования в области культуры, образования, науки. С одной стороны, их представители должны овладеть языком окружающего этноса (что в действительности и наблюдается), с другой – для них важно сохранить язык своих предков (Дуличенко 1998: 26).

One has to note, that the inhabitants of Vershina represent a special case of language islands – a dialect island. The language used by them is a non-standard variety, and exists almost entirely[2] in the oral form, without a stabilized norm and factors supporting the preservation of the language in the minority education system. Such a state favours the intensification of foreign influences (Weinreich 1963: 85-6).

2 Theoretical assumptions

2.1 Language contact and inter-lingual influence

According to Uriel Weinreich's definition "languages are in contact if they are alternately used by the same persons" (Weinreich 1963: 1). This phenomenon belongs to *parole*, which means that the processes associated with it occur on the "fluency continuum" between perfect bilingualism and pure monolingualism (Myers-Scotton 2005: 43). So it is not a state that is achieved once and for all, but a living process and there are some disturbances within it. These are primarily the processes of borrowing and changing the language within the utterance. What is the basic difference between them? The first include both matter (MAT-) borrowings, i.e. "direct replication of morphemes and phonological shapes from a source language" and pattern (PAT-) borrowings, i.e. "re-shaping of language-internal structures, [in which] the formal substance or matter is not imported but is taken from the inherited stock of forms of the recipient or replica language (i.e. the language that is undergoing change)" Matras, Sakel 2007: 829-30). The latter mean switching between two language systems and allow you to indicate the boundaries between them (cf. e.g. Poplack 1988: 220; Sayahi 2014: 81). However, it should be borne in mind that an easy and unambiguous indication of the boundaries between languages is possible in full bilingualism, but often "speakers change their language (A) to approximate what they believe to be the patterns of another language or dialect (B)" (Thomason 2001: 142). The example of relations between the Lesser Poland dialect with the Russian language in Vershina, belongs to such cases where speakers, not having mastered one of the codes fully, make certain approximations by adapting foreign elements and uncontrollably changing the language to maintain conversation.

---

[2] Although there are few texts written in the Lesser Poland dialect in Vershina, they are limited to the lyrics of the folk band Yazhumbek [Less. Pol. Hazel grouse] and have practically no impact on everyday communication and code-switching and mixing phenomena (cf. Ananiewa 2013).

In the case of a change of language during the speech for the purposes of this article, a clarification will be needed, which we will make after Peter Auer, distinguishing code-switching as the juxtaposition of codes "perceived and interpreted as a locally meaningful event by participants" and mixing as those cases, which "cannot be labelled language A or language B", mainly due to the frequency of switches (Auer 1999: 310-4). The diversity of the latter was embraced in the theory of Peter Muysken, who defined the following types of code mixing:

a) Insertional (an element from L_1 language inserted in L_2 sentence; Muysken 2000: 3), e.g. *t'eras na konc'erty xoć stav''ajom to*$_{POL}$ **ńikt'o**$_{RUS}$ *ńe xce iść na k'oncert*[3]$_{POL}$ [Although they are organizing concerts now, **no one**$_{RUS}$ wants to go to the concert]; *c'ałe ž'yće*$_{POL}$ ‖ **d'a**$_{RUS}$, | *ja c'aụe ž'yće p'isoł, z ʒ́eśunt'ygo r'oku i p''iše*$_{POL}$ **pr'ėdisł'ov'ė**$_{RUS}$ *sf'ojej kš'onšk'i žeb'yśće v'eʒ'el'i*$_{POL}$ [Whole life$_{POL}$. **Yes**$_{RUS}$. I was writing whole my life, since I was ten and I am still writing$_{POL}$. **The fore-word**$_{RUS}$ to my book, you know?$_{POL}$]. The main language of the utterance is L_1, and the inserted L_2 element is not subject to integration to L_1, but is transferred in its original form and content. In other words, L_1 serves as the morphosyntactic frame for the clause, and the L_2 elements are "embedded" in it (cf. Myers-Scotton 2004: 106-7).

b) Alternational (full shift from L_1 to L_2 in a single speech act), e.g. *muj ʒ́'adek m'aụ mụyn | gdy zmarụ | bur''aći b'arʒo pụak'al'i*$_{POL}$ ‖ *xar'ošyj čėłav''ek | ńikạgd'a ńė ạtk'azyvəł*$_{RUS}$ ‖ *xoʒ́‿ʒ́eń | xoʒ́‿noc pš'yjm'e | tag‿go fspom'in'al'i*$_{POL}$ [My grand-father had a mill. When he died, Buryats mourned him a lot$_{POL}$. **He was such a good man, he never refused**$_{RUS}$ (to help). It did not matter for him if it was day or night. That's how they remembered him$_{POL}$];

I1[4]: *n'o j'o b'y z'aros porozm'ov'oụ po p'olsku* [Well, I would gladly speak Polish now]

I2: *j'es'l'i v''iće že j'es'l'i co jo ńe tak no ńe t'ak p'ov'im to* [If you notice that I tell something not correctly]

I3: *no no* [to granddaughter] ***fs'o* | *v''ižu* | *b'ol'še ńelz''a***$_{RUS}$ [**That's all. I see you do not need any more**].

I2: *b'o j'o…* [because I…]

I1: *n'e v'y t'o d'obže rozm'ov'oće p'o p'o polsku | b'o t'ak'im j'ag m'y | ńikt'uźy pšyjiʒ́ǯ'ajum b'arʒo ńe 'iʒe poj'uńć t'ak'e sw'ova* [No, you are speaking Polish very well, just like us. Because sometimes those who come here are speaking in the way we do not understand].

According to Muysken, alternation is the only form of juxtaposition of the codes during speech that can be defined not as mixing but as code-switching (Muysken 2000: 4). These types of switches depend on the topic, but also other conditions: place, time, participants of communication (Gumperz 1977: 1). In the cited examples, the speakers change the code either because they are quoting some else's words with an introducing phrase *bur''aći b'arʒo płak'al'i* [Buryats mourned him a lot] or because they

---

[3] All examples have been written in a simplified Slavic transcription, and grammatical descriptions have been prepared according to the Leipzig Glossing Rules.

[4] If there were more participants in the conversation, their utterances are numbered.

are addressing a person who does not speak Polish (a granddaughter). Such changes include a full switch in terms of the language of the matrix as well as the elements filling it from L_1 to L_2.

    c) Congruent lexicalization (largely or completely shared syntax structure, lexicalized by elements of either language), e.g.

> to śe stro'iṇo_MAT <u>tam i tam to to</u> **fstr''ećʹa byw'a fs''ex**_RUS | <u>kto</u> co skunt pšy… popšyj'iž:oṇ i **ugoš'eńije**_RUS <u>b'yṇo fs'o</u>_RUS <u>b'yṇo</u> d'obže | **da**_RUS | no a p'otym jo ńe v'ėm | coś ńe st'al'i tud'a zax'oẑić kr'om'e_RUS kość'oṇu_POL ńikud'a_RUS | a c'yje <u>to to</u> **ńe znaj…**_RUS ńe v'im [It was being built_MAT there, and there **was the meeting for everyone**_RUS | who came here from various places, there was **the treat**_RUS, **everything**_RUS was good. **Yes**_RUS. And after that, I do not know. For some reason **we stopped going there, except**_RUS of the church_POL, **nowhere**_RUS. And whose it is? **I do not kn…**_RUS I do not know].

This is a special type of code mixing in which some elements of the other two can be find, and words from both languages are "inserted more or less randomly" (cf. Muysken 2000: 3-5). In the case of the above excerpt, the intention of the speaker was to formulate a statement in Polish, but there were cases of inserting single words (*ugoś'eńije* [the treat], *fs'o* [everything], *da* [yes]) and component sentences (*fstr''ećʹa byw'a fs''ex* [there was the meeting for everyone]; *ńe st'al'i tud'a zax'oẑić kr'om'e*_RUS *kość'owu*_POL *ńikud'a*_RUS [we stopped going there, except_RUS of the church_POL, nowhere_RUS]), which, according to the interpretation of many researchers, should be treated as alternations (cf. Auer 1999: 309-10; Myusken 2000: 3-4; Deuchar 2020: 1-2). The situation is additionally complicated by the fact that in the second sentence there is a Polish element of *kość'owu* [church_GEN.SG], used in accordance with the dialectal paradigm of inflection, so it cannot be treated as a Polish borrowing in a Russian utterance, but as an insertion.

There is also a switch, which resulted from self-correction: ***ńe znaj…***_RUS ńe v'im_POL [**I do not kn…**_RUS I do not know_POL]. The informant started sentence in Russian, but did not finish it and quickly came back to the main language of the utterance. Russian influence is also visible in a MAT-borrowing *śe stro'iṇo* [it was being built] (cf. Rus. *строилось*), which has been integrated in the target language. Moreover, due to the structural closeness of two Slavic language in contact, the utterance contains words that could be characterized both as Polish and Russian, e.g. *to* [it], *tam* [there], *kto* [who] and other, underlined in the analysed passage.

In order to analyse such cases, after Svyatlana Tesch and Gerd Hentschel, two levels of articulation will be distinguished. The first embraces the lexis and syntactic structures as well as those elements of the expression plan that determine their morphological form in the mind, i.e. they relate to the content plan (*signifié*). The second level is a phonetic-phonological surface, i.e. the plan of expression (*signifiant*). Thus, there are three possible convergence relations, visible primarily in cognate languages: a) both at the deep-morphonological and phonetic-phonological level, e.g. *tam*_POL and *tam*_RUS [there]; b) only at the deep level, while the phonetic-phonological form is different, e.g. *b'yła*_POL and *b'yła*_RUS [was_PST.F.3SG]; c) the units are different on both levels, e.g. *fšystko*_POL and *fs'o*_RUS [everything] (Теш, Хенчель 2009: 210). Of course, if congruently lexicalized words with the same morpho- logical or morpho- phonological and

phonetic structure occur in the course of a longer utterance in one of the languages, we treat them as an element of a given code, e.g. the negation *ńe* [not$_{NEG}$] in the sentence *no a p'otym jo ńe v'ėm* [and then I do not know], despite to the similarity to the second language. However, if they occur at the junction of fragments in L$_1$ and L$_2$, they can be considered part of each of them, e.g. *b'yu̯o* [was$_{PST.N.3SG}$] in *ugoš'eńije*$_{RUS}$ *b'yu̯o* *fs'o*$_{RUS}$ *b'yu̯o d'obže*$_{POL}$ [there was the treat, everything was good]. Since it is impossible to unambiguously attach this word to any of the codes, it is a prototypical example of congruent lexicalisation, and the entire excerpt, due to the mixture of elements both form L$_1$ and L$_2$ can be characterized as such.

2.2 Different language levels and the phenomena of language contact

An attempt to break down congruently lexicalized speech into prime factors is to relate it to the analytical levels of language distinguished in structuralism: phonetics and phonology, morphology, lexis, syntax (Haas 1960: 267). In the traditional structuralist approach, foreign language influences are analysed at different levels of language (cf. e.g. phonetic, grammatical and lexical interference in Weinreich 1963). In considering the nature of these influences, it should be taken into account that all residents of Vershina speak Russian at the level of monolingual native speakers, and the knowledge of the Lesser Poland dialect is varied and ranges from the communicative minimum among the younger generations in mixed families to a relatively high level among older people living in homogeneously Polish families. The Russian language is mastered both in everyday communication and in the course of school education, which results in a strong orientation to the norm and limits the scope of interference to a minimum. This results in an uneven nature of the linguistic influence: changes under the influence of contact are observed practically only in the heritage language, i.e. the Lesser Poland dialect. The switches that take place between the two languages take place as part of utterances intentionally formulated in Polish, in which Russian elements are intertwined, but also in the matrix language itself, the Russian influence is also present, although not to varying degrees on all levels. As Svetlana Mitrenga-Ulitina notes, the phonetic system of the Vershinian dialect is the most resistant to Russian influence (Mitrenga-Ulitina 2015: 145), which is also confirmed by our analyses. The characteristic features of the Vershinian dialect, such as mazuration (or mazurism), i.e. the pronunciation of the alveolar consonants š, č, ž, ǯ as dental s, c, z, ʒ (see Sawicka 2020), have not only been preserved in the original Polish lexicon, e.g. *c'ysty* [clean, pure] cf. stand. Pol. *czysty* and Rus. *чистый*, but also appear in adapted borrowings from the Russian language, e.g. *pol'acka* [a Polish woman$_{COLL}$], cf. Rus. *полячка*.

The Russian influence at the morphological level is largely connected with the lexical one. In the Vershinian Lesser Poland dialect, borrowed morephemes are stems, not prefixes or suffixes, and therefore, according to the most widespread point of view in the study of language contacts (cf. e.g. Poplack, Sankoff 1984: 104; Thomason, Kaufmann 1988: 37), they should be treated as a form of adaptation of borrowings through the exchange of affixes, and not interference at the level of morphemes. The borrowed stems take a phonological form and are equipped with affixes typical of the

recipient language, such as -*stroj*- [-build-] in its basic meaning 'to build': *o te d'umy v''incėj o tak **stroj'une** b'ywy* [Oh, these houses **were being built** in this manner]; *d'um t'ak'i v''e:lg'i **postroj'iwy*** [They **built** such a big house], as well as in the metaphorical colloquial meaning 'to find a place, to be settled': *'one f pos'olstvo kajś **pšystroj'une*** [They **have been settled** somewhere in the embassy]. The combination of lexical and morphological influence is also visible in newly coined words reproducing Russian models in the form of word-formation calques (i.e. PAT-borrowings) morpheme-by-morpheme, e.g. ***pš'ev°ėść*** *na nacjɔn'alny j'iĩʑyk* [to translate into the national language], cf. Rus. *перевести* [to translate] – the word has been created in the course of replacement of Russian morphemes with the Polish ones: *p'er'e-* : *pše-*, -*v'est'i-* : -*v'eść-*.

However, there are cases that can be described as half-calculus or linguistic hybrids, e.g. *p'otem **zapšec'one** b'yu̯o fšo* 'after that everything was forbidden' – the new word *zapš'ecać*$_{\text{INF}}$ was created under the influence of Rus. *запрещать* [to forbid]. The morpheme – *pr'e-* was replaced with its Pol. equivalent -*pše-*, and the second part has been phonetically adapted in the course of sound substitution *š:'/šč'* : *c*. It is a similar case to Polish *bawełna*, or Czech *bavlna* [cotton] reproducing German *Baumwolle*, where the word is partially "translated" (morpheme substitution) and partially phonetically adapted (sound substitution) (cf. Weinreich 1963: 57-63). The new word has Polish conjugation and can be considered a form of PAT-borrowing.

The most common and most noticeable, and thus the best described manifestation of foreign influence at the lexical level are MAT-borrowings, i.e. single words and complex lexical units, which appear in the target language because of the need to describe a fragment of reality that requires the use of vocabulary present only in the source language. The process of borrowing in this respect has a lot in common with the insertional type of code-mixing. The main cause of the use of an integrated (MAT-borrowing) or not integrated (insertion) foreign element is code repair (cf. Gafaranga 2012: 509-10). Since these processes are especially vivid in insular communities experiencing deficit of their lexical resources (cf. Дуличенко 1998: 26), they are frequent in the Lesser Poland dialect in Vershina, too, e.g. *bestau̯kovyj* [dumb, fool], cf. Rus. *бестолковый*: *ń'e sux'ajćė t'ak'ėgo **bestau̯kov'ėgo*** [Do not listen to this fool]; *iz'ucać* [to study$_{\text{INF}}$, to learn$_{\text{INF}}$], cf. Rus. *изучать*: *gr'upa pɔjex'aa v:ak'acᵛji j'iĩʑyk | **iz'ucać** p'olsk'i j'iĩʑyk* [the group went for holidays to study Polish]; *dvojń'ašk'i* [twins$_{\text{PL}}$], cf. Rus. *двойняшки*: *ʒ'ou̯xu u mń'e **dvojń'ašk'i*** [My daughters are twins]. The borrowed items have been adapted to Polish inflectional system, and *iz'ucać* also reflects the phenomenon of mazuration (*č > c*).

A type of interference characteristic at the syntax level are PAT-borrowings including reproduction of syntactic structures (syntactic calques), e.g. constructions expressing time relations, such as location of events in hourly time: *v + X goʑiny/goʑin*$_{\text{NOM}}$ following the Russian scheme *в X часа/часов* [at X o'clock], cf. Pol. *o X godzinie: f śt'yry goʑ'iny ś'e fstav'awɔ | i do v'ėc'ora do p'uźna ś'e rob'iwɔ'* [We were getting up at 4 o'clock and worked until the evening]. Another example is the expression of an event taking place before a specified time: *do + X*$_{\text{GEN}}$ [before X], cf. Pol. *przed X*$_{\text{INS}}$: *ješč'o **do vojny**$_{\text{GEN}}$ |

a ẑ'adek ve v'ojne 'umarṷ [It was before the war, and the grandfather died during the war]. Because PAT-borrowings are not foreign elements, but a reproduction of foreign models using the resources of the recipient language, they are more difficult to notice, and they are often treated as native by speakers of a given language.

At the syntax level, there are also switches described in paragraph 2.1, but the example of congruent lexicalization in section (c) shows that the distinction of insertions and alternations, as well as manifestations of phonetic, morphological and lexical, as well as syntactic interference is possible in short excerpts – single phrases and sentences. In longer utterances, they overlap and mix, causing, for example, that the alternation is made not to pure L_2 code, but to L_2 containing PAT- and MAT-borrowings and insertions. Therefore, we decided that in addition to the traditionally mentioned analytical levels of language: phonetics, morphology, lexis and syntax, the analysis of code-switching and mixing should also take into account the most complex level – text. Such approach is in line with Francois Grosjean's postulate of a "wholistic" view on bilingualism, according to which there are no isolated instances of juxtaposition of codes, because "the bilingual is an integrated whole which cannot easily be decomposed into two separate parts", and the texts he or she produces are a "unique and specific linguistic configuration" (Grosjean 1992: 54-5). Having made this addition, we can summarise the phenomena occurring at successive analytical levels of the language in terms of both borrowings and code-switching/mixing in the table below.

Table 1: Forms of interlingual influence on different language levels.

Language level	Code-switching and code-mixing	Borrowing processes
phonetics and phonology – phonemes	n/a	phonetic interference
morphology – morphemes	n/a	morphological interference, language hybrids
lexis – words	insertions	PAT-borrowings (word-formation and morphological calques), MAT-borrowings („classical" borrowings)
syntax – phrases and sentences	alternation, insertions	PAT-borrowings (syntactic calques)
text	all listed above	all listed above

Since only at the textual level are individual phenomena visible in a broader context showing their mutual connections and dependencies, we will adopt this perspective when discussing examples of this phenomenon in the Vershinians' speech.

3 Discussion

Treating not only the processes of code-switching and mixing, but the entire Vershinian bilingualism as texts, rather than separate sentences, allows us to illustrate the most complete approach to this phenomenon. The first excerpt is an example of a text maintained relatively consistently in the Lesser Poland dialect. The informant – an elderly woman, is answering of the researchers, and is aware of the lexical deficit in the heritage language, but nevertheless she is able to control the main language of the utterance:

> *zad'ėjće m'i j'ak'ė te vapr'osy | po ros'yjsku* [śmiech] *vapr'osy | bo jo | b'abuške*$_{MAT}$ *po p'olsku zv'al'i*$_{MAT}$ [first name] *a po r'usku* [first name] *b'ėńżė dva l'ata | j'agžė śe to naz'yvo | ś'erp'iń | vž'eśiń | b'ėżė u̯ośėmnast'ėgo vž'eśńa dva l'ata jag' już ńė v''iʒe ń'ic | fcale* [Please ask me some <u>questions</u>. In Russian (laughter) – <u>questions</u>. Because I… My grandmother's$_{MAT}$ na-me$_{MAT}$ in Polish was (first name) and in Russian (first name). It will have been… How was it? August? September? By September 18, it will have been two years since I do not see antyhing. At all].

> - Wcale? [*At all?*]

> *i gu̯'uxom | na j'ed… na to 'uxo jo juš to no dvaj'eśća* ᵘ*'ośėm lot ńė su̯'yše | a na to žė tak tr'oške su̯'yšė | d'obžė <u>vr'oʒė</u> žė rozm'ov'o al'e co | ńic ńė ńė pojm'ujė* [And I am deaf. Out of one… I have not been able to hear out of one ear for 28 years now. And with this other ear I can only hear a little: someone <u>seems to be</u>$_{PRT}$ talking something, but what? I do not understand anything].

> - A do doktora pani jeździła? [*And did you go to the doctor?*]

> *no mńe b'yu̯o v: 'ośėmʒ'eśunt t'yśu̯ʒ ʒev''ińćėt ośėmʒeś'uu̯tym r'oku d'v'e opėr'acjė na gu̯'ove | ᵘot | i m'e z'aro sparal'ižov'au̯o na str'one | jak m'i 'ino zėrv'al'i tėn nu | po ros'yjsku tr'ojńič-nyj ńerf nap'is'al'i | nu pojmuj'ećė mńe* [Well, in 1988 I had two head operations. Yes. And it immediately paralysed my side. As soon as they ruptured this… in Russian it is the <u>trigemi-nal nerve</u>. <u>As they wrote</u>. So do you understand me?]

There are only two unambiguous cases of a MAT-borrowing: *b'abuška* [grand-mother], which has been adapted to Polish declension (cf. Pol. *b'abuške*$_{ACC}$ and Rus. *бабушку*$_{ACC}$), and *zv'al'i* [they called$_{PST.3PL}$] (cf. Pol. *naz'yvać* and Rus. *звать*). The other instances of possible MAT- and PAT-borrowings or code switching and mixing has been underlined, because their interpretation is not obvious:

- The form of *vapr'osy* [questions] in NOM.PL would be the same both in the case of Russian original and its Polish adaptation. Thus, on the basis of the given excerpt is impossible to answer this question and the reference to the entire corpora is needed to check the frequency and other forms. There are only 5 instances of the use of this word and all of them in NOM.SG and NOM.PL, which do not show any differences in inflection. However, the Polish equivalent *pytanie* is much more frequent and

productive. The words built with a help of the stem -*pyt*-, nouns and verbs, e.g. *pytać*, *spytać*, *zapytać* [to ask$_{IMP/PERF}$] appear 23 times, which suggests that *vapr'os* has not been fully established in the Vershinian dialect yet and should be considered either an insertions or a nonce borrowing (cf. Halmari 1997: 17).

- The first part of the last underlined phrase *tr'ojńičnyj ńerf* [trigeminal nerve] could be treated as a prototypical example of insertional code-mixing: due to lexical deficit an item from L$_2$ is embedded in the L$_1$ statement. However, the speaker first hesitated looking for the right word, and then announced the change of language, therefore it was conscious, as in the case of alternation. The decisive factor in such cases is the continuation of the utterance: if there is a return to the original language of the utterance (L$_1$), it is an insertion, and if the utterance is continued in L$_2$, one should classify it as an alternational switch. The continuation of this statement, however, is not unambiguous, because the form *nap'is'al'i* [they wrote$_{PST.3PL}$] is identical in L$_1$ and L$_2$, both at the morphological and surface-articulation level (cf. Теш, Хенчель 2009: 210). Since the rest of the utterance is in Polish, *nap'is'al'i* is a congruently lexicalized element in the common syntactic structure, and may fulfill the function a switch point between L$_1$ and L$_2$ in this fragment of utterance: *zěrv'al'i těn nu | po ros'yjsku*$_{POL}$ *tr'ojńičnyj ńerf*$_{RUS}$ **nap'is'al'i**$_{RUS-POL}$ | *nu pojmuj'ećé mńe*$_{POL}$ [they ruptured this… in Russian it is$_{POL}$ the trigeminal nerve$_{RUS}$. **As they wrote**$_{RUS-POL}$. So do you understand me?$_{POL}$].

The analysed example shows that although interpretations of text excerpts may lead to unambiguous classifications of individual words, word combinations and phrases, in a broader context their ambiguity is revealed. The interpretations in the next fragment are also ambiguous, where, however, the frequency of clear L$_1$–L$_2$ switches and Russian influences is higher.

jak j'uš st'aľin 'umar to to | i fs'ʼo p'ošu̯o$_{PAT}$ | **d'emokr'at'ija**RUS dr'ugo st'au̯a$_{PAT}$ | a t'ak demokr'ac'ji ńe b'yu̯o | f kou̯x'oźe | f proizv'octfax | **na f'abr'ikax** | **rukovod'ʼit'el'i b'yl'i**$_{RUS}$ | 'uny m'ʼau̯y t'ʼekst | t'yle pov'ińːiśće zar'ob'ać | t'yleśće d'ou̯žńi$_{MAT}$ zapu̯'aćić l'u̯žum | t'yle d'ou̯žno$_{MAT}$ iś gosud'arstfu$_{MAT}$ | uk'azane$_{MAT}$ | a t'eras **ńikt'o**$_{RUS}$ ńik'omu ńe uk'azuje$_{MAT}$ | k'aždy targ'uje | jak p'ošu̯a ta | sfob'odno targ'ovl'a$_{MAT}$ | to | fs'ʼo pošu̯'o po in'akšymu$_{PAT}$ | no co zr'ob'iće | no | **nać'alstfo j'est nać'alstfo**$_{RUS}$ | v'ʼiźi na očax$_{PAT}$ že k'ʼepsko r'ob'i | j'ego spr'ava | i k'ʼepsko | j'esl'i s šer'egu śe od'ezv'eš | to ty j'uš **vr'ak nar'oda**$_{RUS}$ | pšysf'ajivaľi t'ak

[Once Stalin died, everything went on$_{PAT}$. Another **democracy**$_{RUS}$ began$_{PAT}$. And before that, there was no democracy in kolkhoz, in production, **in factories**$_{RUS}$. There were directors. And they had an instruction: you are supposed to earn this much, this much you have to$_{MAT}$ pay the people, this much you have to$_{MAT}$ pay to the state$_{MAT}$. (It was) specified$_{MAT}$. And now **no one**$_{RUS}$ is giving orders$_{MAT}$ to anyone. Everyone is trading. When the free market economy$_{MAT}$ began, everything went on differently$_{PAT}$. What can you do? **The directors are the directors**$_{RUS}$. They see with their own eyes$_{PAT}$, that someone is working badly. It is his business. And it is not good when you make any comment – you are already **the enemy of the people**$_{RUS}$. (Such a man) was described this way].

There are several instances of possible insertions, but only two of them are not combined with another type of code-mixing: *ńikt'o* [no one] cf. Rus. *никто* and *vr'ak nar'oda*

[the enemy of the people]. There is also *d'emokr'at'ija*$_{NOM.SG.RUS}$ and *demokr'ac'ji*[5]$_{GEN.SG.POL}$, but the interpretation is not clear, because a) first instance is Russian and the second one is Polish, b) they appear in a longer passage of PAT-borrowings (*fs''o p'ošu̯o* [everything went on] cf. Rus. *всё пошло* and *dr'ugo st'au̯a* [became different] cf. Rus. *другая стала*) combined with insertion and is switched to a fragment of utterance that can be interpreted both as Russian and as MAT-borrowings (*f kou̯x'oźe | f proizv'octfax* [in kolkhoz, in produc-tion]). The part i *fs''o p'ošu̯o*$_{PAT}$ | *d'emokr'at'ija*$_{RUS}$ *dr'ugo st'au̯a*$_{PAT}$ | a *t'ak demokr'ac'ji ńe b'yu̯o | f kou̯x'oźe | f proizv'octfax | na f'abr'ikax | rukovod'it'el'i b'yl'i*$_{RUS}$ is predominantly Russian, but because the presence of elements integrated to Lesser Poland dialect and shared lexis[6] determines the excerpt in whole as congruent lexicalisation. The speaker's intention was to speak Polish, but he was not aware of the switch points.

The excerpt **nać'alstfo j'est nać'alstfo**$_{RUS}$ [The directors are the directors] is a separate sentence, which could be treated as a full switch from L$_1$ to L$_2$, i.e. alternation, but as a fixed phrasematic structure with limited possibilities of the changes in the scheme *X есть X* [*x is x*], has a lot in common with insertions. Muysken pointed out that the types of code mixing distinguished by him are not independent states, but that real phenomena are located on continuums between the three ideal types. The presented example shows just such an intermediate state between insertion and alternation.

There are frequent MAT-borrowings. Those related to the domain of economics or politics can be characterised as cultural borrowings, i.e. names for objects and process-es acquired in the course of socio-cultural contacts, for which the recipient language had no names before (cf. Sayahi 2014: 89; cf. Weinreich 1963: 53-4): *gosud'arstfo* [state] Rus. *государство*, *sfob'odno targ'ovl'a* [free trade, free market economy] Rus. *свободная торговля*. However, the others are replacing words that already existed in the recipient language (Myers-Scotton 2005: 215): *d'ou̯žno*, *d'ou̯žńi* [should] Rus. *должно*, *должны*, *uk'azane* [specified] Rus. *указано*.

The prepositional phrases of phrasematic character are clear examples of PAT-borrowings: *po in'' akšuti* [in a different way, differently] reproduces Rus. *по-другому*, and *na očax* [with someone's own eyes] – Rus. *на глазах*. The last highlighted item – *pšysf'ajivać* was used under the influence of Rus. *присваивать* [assign$_{INF}$] and may be treated both as a PAT-borrowing (word-formation semi-calque in which the prefix is replaced with its Polish equivalent and the stem is borrowed), and as an adopted borrowing (MAT-borrowing), in which the replacement of prefix is a form of morpho-nological adaptation (cf. Grek-Pabisowa 1999: 225-6).

Despite the intentions of the speaker, who was an elderly man with a good command of both languages (with Russian as the language better known and used more frequently

---

[5] Polish *demokracja* and Russian *демократия* are an example of words identical only on the morphono-logical level, but the differences on the phonetic level refer to *d : d'* and *c'ja : t'ija*. Such similarities favour mixing (congruent lexicalisation) (cf. Muysken 2000: 1-5).

[6] *D'emokrac'ija*, *kou̯xoz*, *proizvoctfo*, *rukovod'it'el* are of Russian origin, but appear in the corpus as adopted borrowings.

in recent years), it was not possible to produce the entire utterance in the Lesser Poland dialect. In his efforts, our interlocutor focuses on a version of the heritage language that he considers native, i.e. containing numerous lexical and syntactic influences of the Russian language (cf. Mitrenga-Ulitina 2015: 130, 146-7).

4 Conclusion

In the fragments discussed, the transition between different types of code switching and mixing as well as forms of borrowing often takes place depending on the adopted perspective: partial or "wholistic". The analysis of the Vershinians' bilingual statements on the text level shows that many phenomena, as in the case of congruent lexicalization in dysfluent speech defined by John Lipski, result from: incomplete fluency in one of the languages coupled with the intention to maintain the utterance in this language, as well as from the lack of social consequences for involuntary mixing (Lipski 2009: 33). Attempts to complete or fix the code by our informants are too visible to talk about a homogeneous L_1 language that has so many features of a mixed code (or a fused lect, cf. Auer 1999: 309-10). Among the types of code-mixing, it is congruent lexicalization, and especially its variant referring to dysfluent speech, that is the most adequate to describe the processes taking place as part of the contact of cognate languages. However, it is only in the analysis at the text level that all dimensions of the interrelationships of various types of juxtaposition of the codes and borrowing processes are fully visible.

References

Natalia Ananiewa, 2013: Teksty Polaków pisane grażdanką w syberyjskiej wsi Wierszyna. *Acta Baltico-Slavica* 37. 287-98.

Peter Auer, 1999: From code-switching via language mixing to fused lects. Toward a dynamic typology of bilingual speech. *The International Journal of Bilingualism* 3. 309-32.

Ludwik Bazylow, 1975: *Syberia*. Warszawa: Wiedza Powszechna.

Barbara Bullock, Almeida Toribio, 2009: Themes in the study of code switching. *The Cambridge handbook of linguistic code-switching*. Eds. Barbara Bullock, Almeida Toribio. Cambridge: Cambridge University Press. 1-17.

Margaret Deuchar, 2020: Code-Switching in Linguistics: A Position Paper. *Languages* 5/22. 1-19. Also online.

Thomas Eriksen, 1993: In which sense do cultural islands exist? *Social Anthropology* 1/1B. 133-47. Also online.

Joseph Gafaranga, 2012: Language alternation and conversational repair in bilingual conversation. *International Journal of Bilingualism* 16/4. 501-27. Also online.

Michał Głuszkowski. 2012. Bilingwizm z dyglosją w sytuacji wyspowej. Polska wieś Wierszyna na Syberii. *Z polskich studiów slawistycznych. Seria 12. Językoznawstwo. Prace na XV Międzynarodowy Kongres Slawistów w Mińsku 2013*. Ed. Małgorzata Korytkowska et al. Warszawa: Polska Akademia Nauk. 45-51.

Michał Głuszkowski, 2015: Creolization and balkanization as a result of language (dialect) contact. Is the origin of mixed languages universal? *Slavia Meridionalis* 15. 53-66. Also online.

Iryda Grek-Pabisowa, 1999 (1983): Typy zapożyczeń i sposoby przyswajania wyrazów polskich. *Staroobrzędowcy. Szkice z historii, języka, obyczajów*. Warszawa: SOW. 211-27.

François Grosjean, 1992: Another view on bilingualism. *Advances in Psychology* 83. 51-62. Also online.

John Gumperz, 1977: The Sociolinguistic Significance of Conversational Code-Switching. *RELC Journal* 8/2. 1-34.

William Haas, 1960: Linguistic Structures. *Word* 16/2. 251-76.

Helena Halmari, 1997: *Government and codeswitching: Explaining American Finnish*. Amsterdam, Philadelphia: John Benjamins Publishing Company.

John Lipski, 2009: "Fluent dysfluency" as Congruent Lexicalization: A Special Case of Radical Code-Mixing. *Journal of Language Contact* 2/2. 1-39. Online.

Heinrich Löffler, 1987: Sprache und Gesellschaft in der Geschichte der vorstrukturalistischen Sprachwissenschaft / Language and Society in the History of Prestructuralist Linguistics). *Sociolinguistics. An International Handbook of the Science of Language and Society. Soziolinguistik. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft. Vol. 1*. Eds. Ulrich Ammon, Norbert Dittmar, Klaus Mattheier. Berlin, New York: Walter de Gruyter. 379-89.

Yaron Matras, Jeanette Sakel, 2007: Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31/4. 829-65. Also online.

Swietłana Mitrenga-Ulitina, 2015: *Język polski mieszkańców wsi Wierszyna na Syberii*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.

Peter Muysken, 2000: *Bilingual speech. A typology of code mixing*. Cambridge: Cambridge University Press.

Carol Myers-Scotton, 2004: Precision Tuning of the Matrix Language Frame (MLF) Model of Codeswitching. *Sociolinguistica* 18. 106-17.

Carol Myers-Scotton, 2005: *Multiple voices. An introduction to bilingualism*. Malden: Blackwell.

Ewa Nowicka, 2011: *Wierszyna, czyli z morza na ocean*. Paper at the conference "Słowiańskie wyspy językowe i kulturowe". Nicolaus Copernicus Univesity in Toruń, September 16-17th 2011.

Shana Poplack, 1988: Contrasting patterns of code-switching in two communities. *Codeswitching: anthropological and sociolinguistic perspectives*. Ed. Monica Heller. Berlin-New York, Amsterdam: Mouton de Gruyter. 215-44.

Shana Poplack, David Sankoff, 1984: Borrowing: the synchrony of integration. *Linguistics* 22. 99-135.

Paul Rosenberg, 2005: Dialect convergence in the German language islands (Sprachinseln). *Dialect change: convergence and divergence in European languages*. Peter Auer, Frans Hinskens, Paul Kerswill. Cambridge: Cambridge University Press. 221-35.

Lofti Sayahi, 2014: *Diglossia and Language Contact. Language Variation and Change in North Africa*. New York: Cambridge University Press.

Irena Sawicka, 2020: Merger of Alveolar and Dental Fricatives and Affricates (Mazurism). *Encyclopedia of Slavic Languages and Linguistics Online*. Ed. Marc Greenberg. Leiden: Brill. Online.

Sarah Thomason, 2001: *Language contact*. Washington: Georgetown Univerity Press.

Sarah Thomason, Terrence Kaufmann, 1988: *Language Contact, Creolization, and Genetic Linguistics*. Berkeley-Los Angeles-Oxford: University of California Press.

Uriel Weinreich, 1963: *Languages in contact. Findings and problems*. The Hague: Mouton.

Александр Дуличенко, 1998: Языки малых этнических групп: статус, развитие, проблемы выживания. *Языки малые и большие... In memoriam acad. Nikita L. Tolstoi*. Ed. Александр Дуличенко. *Slavica Tartuensia* IV. Tartu: Universitas Tartuensis. 26-36.

[Aleksandr Duličenko, 1998: Jazyki malyh etničeskih grupp: status, razvitie, problemy vyživanija. *Jazyki malye i bol'šie... In memoriam acad. Nikita L. Tolstoi*. Ed. Aleksandr Duličenko. *Slavica Tartuensia* IV. Tartu: Universitas Tartuensis. 26-36].

Святлана Теш, Герд Хенчель, 2009: Переключение кодов в трасянке (некоторые количественные наблюдения). *Славянские языки. Аспекты исследования*. Eds. Наталья Ивашина, Елена Руденко. Минск: БГУ. 209-15.

[Svjatlana Teš, Gerd Henčel', 2009: Pereključenie kodov v trasjanke (nekotorye količestvennye nabljudenija). *Slavjanskie jazyki. Aspekty issledovanija*. Eds. Natal'ja Ivašina, Elena Rudenko. Minsk: BGU. 209-15].

Povzetek

Vas Veršina v regiji Irkutsk v Sibiriji so v začetku 20. stoletja ustanovili prostovoljni naseljenci iz južne Poljske. Od začetka je bila dvojezična skupnost, vendar je zaradi družbenih in političnih sprememb ruščina postopoma nadomestila malopoljsko narečje v številnih funkcijah. Dvojezičnost potomcev poljskih naseljencev poleg diglosije spremlja tudi pojav jezikovnih sprememb med govorom. Namen tega članka je obravnavati raznolikost pojavov kodnega preklapljanja in mešanja na različnih jezikovnih ravneh. Na področjih fonetike, morfologije in leksike se raziskovalci najpogosteje osredotočajo na medjezikovne vplive v obliki interference, medtem ko je kodno preklapljanje pojav, ki se pojavlja na ravni skladnje.

Članek temelji na jezikovnem gradivu, zbranem med terenskimi odpravami v Veršini. Opredeljuje primere leksikalnih in strukturnih izposojenk ter različne oblike jezikovnega preklapljanja med govorom. Te so v literaturi tradicionalno analizirane kot: insercije, alternacije in kongruentna leksikalizacija. Vendar pa tak pristop omogoča opisovanje le izbranih delov izreka, medtem ko v živem sporazumevanju delujejo celotna besedila. Zato je bilo predlagano, da se obravnava najkompleksnejša raven jezika, tj. besedilo kot tako. Če se analizira celotno besedilo, se v njem prepletajo različne vrste kodnega preklapljanja in interference, in tisto, kar bi bilo na ravni posameznega stavka opredeljeno kot izposojanje ali alternacija, se v resnici izkaže kot kongruentna leksikalizacija. Izbrani primeri kažejo, kako to dopolnjevanje vpliva na dojemanje pojavov dvojezičnosti.