

The Polish-Russian Mixed Code in the Polish Community in Lithuania



Brigita Séguis

Hertford College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2014

The data for the present study come from a corpus of spontaneous conversations involving members of the Polish community. The recordings were collected in the city of Vilnius and feature 25 respondents in their twenties. The data analysis reveals that all three types of language alternation feature in the present corpus; however, code-mixing is clearly the preferred type. It immediately manifests itself in the frequent insertion of Russian single switches and larger constituents into the Polish base. As a result of its wide spread and frequency, language alternation has lost its immediate local meaning and the pattern of frequent Polish-Russian mixing has become the accepted mode of speaking within the Polish speech community.

Acknowledgements

This thesis would have been impossible without the help of many people, to whom I would like to express my heartfelt thanks. My gratitude goes first and foremost to my parents, who made my time at Oxford possible, and to my husband Jean-Charles who encouraged me to apply to Oxford in the first place and supported me throughout my studies.

My thanks also go to my supervisor Dr Jan Fellerer (Wolfson College) for his excellent guidance, patience and invaluable comments. I have been extremely lucky to have a supervisor who cared so much about my work, diligently went through countless drafts and re-writes, and who was always generous with his time and ideas. I also thank Dr Rosalind Temple (New College) for the feedback on my transfer of status and confirmation chapters, which was very useful and influential in the way the thesis evolved, as well as the invaluable comments that I received during the viva voce examination. I am also extremely grateful to my external examiner, Prof Dr Ad Backus (Tilburg), who suggested very useful corrections and amendments to an earlier version of this thesis.

A word of thanks also goes to my fellow postgraduate students in the Oxford linguistics department. Dr Damien Mooney deserves extra thanks for kindly helping me with PRAAT. I would also like to thank my husband for taking care of all the technical aspects of my thesis, especially for introducing me to \LaTeX and patiently guiding me through all of its intricacies. I have never felt so fortunate to be married to a mathematician!

I would also like to extend my thanks to all the informants who took part in this study, and who willingly sacrificed their time for my benefit. Without them, nothing of substance would have transpired from this thesis. I am extremely grateful to Dariusz Jankowski and Edmund Brazis for their help with recruiting the informants and collecting the recordings on my behalf.

The funding for my studies came from the M.B. Grabowski Fund Postgraduate Scholarship and Mary Starun Scholarship in Polish Studies, awarded by Hertford College. My heartfelt gratitude goes to both of these awarding bodies. In particular, I would like to thank Mr Wojciech Tobiasiewicz from the M.B. Grabowski Fund, as well as his wife Joanna, for their continued support, friendship and much needed help during the most difficult and trying times. Finally, I would also like to express my gratitude to everyone at the Polish Catholic Mission in England and Wales, especially Monsignor Stefan Wylężek and Reverend Krzysztof Tyliczszak, as well as all the trustees of the M.B. Grabowski Fund for the confidence that they placed in me and my research project.

Contents

1	Introduction	1
1.1	Historical Overview	5
1.2	Linguistic Situation	6
1.2.1	Russification	7
1.2.2	Lithuanisation	8
1.2.3	Linguistic Repertoires	9
1.3	The Regional Variety of Polish	13
1.4	The Regional Variety of Russian	17
2	Theoretical Background	20
2.1	Code-Switching	21
2.1.1	Grammatical Approach	22
2.1.2	Socio-Functional Approach	25
2.2	Code-Mixing and Emerging Mixed Codes	30
2.3	Fusion/Fused Lects	32
2.4	Matrix Language Frame Model and Composite Code-Switching	34
3	Research Questions, Data and Methodology	39
3.1	Hypotheses	39
3.2	Methodological Framework	43
3.2.1	Speech Community	43

3.2.2	Social Network	48
3.3	Data Collection	50
3.3.1	Informants	51
3.3.2	Recordings of Natural Language	53
3.4	Transcription Technique	53
4	Discourse Markers	57
4.1	The Category of Discourse Markers	58
4.2	Discourse Markers and Code-Switching/Mixing	62
4.3	Polish and Russian Discourse Markers in the Present Corpus	66
4.4	The Discourse Particles ‘nu’ and ‘vot’	70
4.4.1	Discourse Particle ‘nu’	73
4.4.2	Discourse Particle ‘vot’	81
4.5	Discourse Markers ‘koroče’, ‘tipa’ and ‘prosto’	87
4.5.1	‘Koroče’	88
4.5.2	‘Tipa’	94
4.5.3	‘Prosto’	99
4.6	Swearwords ‘blin’ and ‘bljad”	103
4.7	Conclusions	111
5	Intraclausal Switches	115
5.1	Switches or Borrowings?	117
5.1.1	Lexical Borrowings/Loanwords in the Present Corpus	122
5.2	Matrix Language Frame Model and the 4-M Model	126
5.3	Single Switches in the Present Corpus	132
5.3.1	Nouns	135
5.3.2	Verbs	146

5.3.3	Adjectives	149
5.3.4	Conclusions about Single Switches	152
5.4	Embedded Language Islands	154
5.4.1	Noun Phrases	156
5.4.2	Prepositional Phrases	158
5.4.3	Verb Phrases	160
5.4.4	Internal Embedded Language Islands	167
5.4.5	Conclusions about ELIs	168
5.5	Structural Borrowing	170
5.5.1	Semantic Borrowing	170
5.5.2	Calques	173
5.6	Conclusion	185
6	Interclausal Switches	188
6.1	Interclausal Switching and ‘Classic’ CS	190
6.2	Discourse-Related Functions of Interclausal Switching	194
6.3	Interclausal Switching within a Turn	197
6.3.1	Switching between Main Clauses	198
6.3.2	Switches between a Main and Subordinate Clause	205
6.4	Interclausal Switches between Turns	209
6.5	Russian Interclausal Switches with Polish Lexical Elements	219
6.6	Conclusions	224
7	Discussion and Conclusions	228
7.1	Prototypical Code-Switching or Code Mixing?	228
7.2	Typological Proximity and its Implications	236
7.3	From Code Mixing to a Fused Lect	239
7.4	Drawbacks of the Present Study and Directions for Future Research	242

7.5	Conclusions	244
	Bibliography	247
	Appendices	
A	Data Collection Forms	262
A.1	Letter to participants	263
A.2	Consent form	265
A.3	Questionnaire	266
B	Transcription Extract	267

List of Tables

1.1	Selection of important features of regional Polish	16
4.1	Russian ‘nu’, ‘vot’, ‘nu vot’ and Polish ‘no’: functions and numbers	72
5.1	Distribution of ‘maszyna/samochód’ in the Polish base	124
5.2	Breakdown of single switched nouns for each case/gender	137
7.1	Overview of possible outcomes of language contact	229

List of Abbreviations

The following is a list of abbreviations used in this thesis, except for the conventionalised ones, such as *e.g.*, *i.e.*, *etc.*

CS	code-switching
CM	code-mixing
MLF	Matrix Language Frame
ELI	Embedded Language Island
NP	noun phrase
VP	verb phrase
PP	prepositional phrase
PL	Polish
RUS	Russian
LT	Lithuanian
colloq.	colloquial
lit.	literally
emph.	emphatic

List of Transcription Symbols

The transcription symbols below and their explanation are based on the system developed by Atkinson and Heritage (1984: ix-xvi).

SYMBOL	SIGNIFICANCE
(0.5)	The number in brackets indicates a time gap in tenths of a second.
(.)	A dot enclosed in a bracket indicates pause in the talk less than two tenths of a second.
(())	A description enclosed in a double bracket indicates a non-verbal activity. For example ((laughter)).
()	Empty parentheses/brackets indicate the presence of an unclear fragment on the tape. The length of the brackets indicates the length of the untranscribed talk.
(guess)	The words within a single bracket indicate the transcriber's best guess at an unclear fragment.
[]	Square brackets between adjacent lines of concurrent speech indicate the onset (and end) of a spate of overlapping talk.
=	Equals sign at the end of one line and in the beginning of next line indicates that there is no 'gap' between the two lines.
:::	Colons indicate prolongation of the immediately prior sound.
—	A dash indicates the sharp cut off of a prior word or sound.

Chapter 1

Introduction

The linguistic situation in the Polish community in Lithuania has, over the years, been studied by focusing on the phonological, morpho-syntactic and lexical features of the variety of Polish known as ‘polszczyzna kresowa’, or ‘borderline Polish’, spoken in the city of Vilnius and the Vilnius region. Starting with the earliest studies (Nitsch 1925; Turska 1939/1982), the main task of the researchers has been to establish the degree of diversion from Standard Polish following the separation of the Vilnius region from mainland Poland. While these studies, discussed in greater detail in section 1.3, offer a valuable insight into the development of the language spoken on the territory of the former Polish borderlands, they provide very little information about the extent of the influence exerted by other languages which this variety of Polish has been in regular contact with.

The present study aims to further expand our understanding of the outcome of prolonged language contact in the Polish community in Lithuania and to evaluate its extent on the basis of spoken language data collected *in situ*, presented in Chapter 3. As will be explained in the remainder of this chapter, several socio-political developments, most notably Russification during Soviet times and Lithuanisation after independence, led to the development of complex linguistic repertoires among the

Lithuanian Poles, which in most cases consist of the regional variety of Polish, Russian and Lithuanian. One of the outcomes of prolonged language contact has been the emergence of code-switching, which in its most general sense can be defined as the alternating use of two or more languages within one conversational episode (Auer 1998: 1). As will be demonstrated in Chapters 4 and 5, in the process of transcribing collected recordings it had already emerged that the data display numerous and frequent cases of language alternation between the regional variety of Polish and Russian, and to much lesser extent Lithuanian. This pattern is in fact reminiscent of what has been referred to by some researchers as ‘code-mixing’, and led to the main research question being reformulated as “can the frequent and regular language alternation in the Polish community in Lithuania be classified as a case of an emerging Polish-Russian mixed code?” The aim of the present study is therefore twofold. The first aim is to evaluate the evidence for the emergence of a Polish-Russian mixed code (Chapters 4 and 5), while the second aim is to assess the evidence for the ‘classic’ model of code-switching (Chapter 6) in order to establish to what extent the two patterns of language alternation co-exist in the Polish community in Lithuania (Chapter 7).

The corpus of transcribed recordings consists of 109,722 orthographic words, out of which there are 89,015 Polish words, 20,093 Russian words and 614 Lithuanian words. It should be noted at the very outset that this quantification has two major drawbacks, and should therefore be treated as an estimate, rather than a precise figure. The first drawback is that there are 155 instances of single-word calques and 83 instances of structural calques, studied in Chapter 5, which can be defined as complex lexical units that are created by a morpheme-by-morpheme or word-for-word translation of the source unit (Haspelmath 2009: 39). Since calques consist of Polish morphological and lexical elements, they have been automatically counted as Polish words despite their underlying Russian structure.

The second drawback stems from the fact that Polish and Russian are two typologically closely related languages, which share a great number of morphemes, both free and bound. While this in itself facilitates switching between the two languages, in some instances it also creates ambiguity in determining whether a given lexical item should be labelled as ‘Polish’ or ‘Russian’. Overall, there are 637 instances of ambiguous lexical items in the present corpus. The majority of them appear at the boundaries between stretches of Polish discourse and Russian switches, e.g. ‘To on /он прилип jak банный лист’ – *It’s he who stuck like a limpet*. The pronoun ‘on’ – *he* is the same in Polish and Russian, which makes it difficult to establish whether in this example the switch starts before or after the pronoun. While in certain cases pronunciation can be employed as the disambiguating cue, there is some evidence of phonetic convergence between the regional Polish and Russian systems, which means that this criterion cannot be applied in every instance and that some lexical items will inevitably remain ambiguous.

Although not completely satisfactory, the quantification presented above reveals several important findings. First of all, it further corroborates the fact that the pattern of language alternation analysed in this thesis is Polish-based. Apart from 60 stretches when the base language is clearly Russian and 38 instances when it is ambiguous, Polish serves as the base language in all the remaining instances, hence it is not surprising that Polish lexical items make up the biggest proportion of the corpus.

Secondly, the quantification given here also reveals the scarcity of Lithuanian lexical items, which are nowhere near as numerous as Russian words. Unlike intraclausal Russian switches, which, apart from discourse markers and swearwords, tend to be idiosyncratic, semantically-varied and non-recurrent, Lithuanian lexical items tend to fall into particular semantic categories. The most prominent category is that of institutional and legal terminology, e.g.: ‘draudimas’ – *insurance*, ‘seniūnija’ – *municipality*, ‘sąjūnga’ – *union*, ‘vadybininkas’ – *manager*. The second most prominent

category are proper nouns, such as names of films, streets or newspapers, e.g.: ‘Pukuotukas’ – *Winnie-the-Pooh*, ‘Greiti ir įsiūtę’ – *Fast and Furious*. The fact that, as opposed to Russian switches, the majority of Lithuanian switches can be classified into specific semantic categories suggests that there must be a semantic motivation, which prompts speakers to use exactly these lexical elements, and not others. This motivation can be neatly explained with Backus’ (2001) notion of ‘semantic specificity’. According to Backus (2001: 134), proper names, institutional terminology and various cultural references designate unique and specific referents. If a speaker has experience with a particular domain in a given language, then he or she may choose to insert the specific lexical item associated with that domain, along with its connotations, rather than look for possible translation equivalents in the receiving language.

Third, the quantification presented above also reveals the extent to which Russian features in the present data. Language alternation involving Russian lexical items is both quantitatively and qualitatively different from the few examples of language alternation involving Lithuanian cited above. Overall, there are 9,111 switches into Russian in the present corpus, which consist of 5,268 switches involving discourse markers, which will be addressed in greater detail in Chapter 4, 3,521 intraclausal switches (single and multiple), analysed in Chapter 5, and 322 interclausal switches, studied in Chapter 6. In the course of the data analysis, it will be argued that typological similarity between Polish and Russian on the one hand, and the prolonged contact between the two languages on the other hand have led to the emergence of several language contact phenomena, namely ‘classic’ or prototypical code-switching, followed by the emergence of Polish-Russian code-mixing, with some potential evidence of a further development into a fused lect. As will be explained in the literature review in Chapter 2, each type of switching is said to be characteristic of a different kind of language alternation phenomenon: ‘classic’ or prototypical code-switching is normally defined by the presence of interclausal and intraclausal switches, frequent insertions of intraclausal switches are one of the characteristic features of code-mixing,

while significant presence of switched discourse markers may be a sign of an emerging fused lect. The correlations between the type of switching and the kind of language alternation phenomenon that it is said to signify will be discussed with regard to the Polish-Russian data throughout the remainder of this thesis.

The aim of the first chapter is to present the historical and social background, which gave rise to the phenomenon of Polish-Russian language alternation, investigated in this study. It is followed by an overview of previous research on the phonological, lexical and grammatical features of the northern ‘borderline Polish’, as well as some brief information about the variety of Russian spoken in Lithuania.

1.1 Historical Overview

The Polish community can be described as an autochthonous ethnic group that has been living on the territory of modern-day Lithuania since as far back as the fourteenth century (Potašenko 2007: 13). Historically, the number of Poles varied during different periods. The first significant influx of the Polish population to the Grand Duchy of Lithuania was recorded between 1569 and 1795, which coincides with the period of the Polish-Lithuanian Commonwealth (Krzywicki 2005: 91). During the Commonwealth period (1569–1795), much of the Lithuanian nobility was polonised and formally joined the prestigious ‘szlachta’ class, which resulted in a further increase in the Polish speaking population (Kurcz 2005: 77).

The Polish community continued to grow steadily throughout the nineteenth and first half of the twentieth century. However, as a result of the change of borders after World War II, a great number of Poles moved back to Poland, and the ones who remained in Lithuania became an ethnic minority in a foreign country (Krzywicki 2005: 94). As Hogan-Brun et al. (2005: 354) explain, the Polish community today consists of “either descendants of ethnic Lithuanian families who were Polonised during the long-

term union with Poland or later [...], or else they are from Polish families who have been resident for generations.” The point to be made here is that, notwithstanding the origins, the Poles represent a historically rooted indigenous ethnic group with a long-established presence on the territory of modern-day Lithuania.

Concentrated in the city of Vilnius, as well as the Vilnius and Šalčininkai regions, the Polish community today makes up the largest ethnic minority in the country, which, according to the Population and Housing Census (2011), consists of 210,600 inhabitants, or 6,6% of the entire population. The majority of Poles are centred in two districts: the Šalčininkai district, where they constitute 79,5 % of the total population, the Vilnius district, where they constitute 61,3%, and the city of Vilnius where the number of Poles amounts to 19,7%.

1.2 Linguistic Situation

The linguistic situation in the Polish community is rather complex, which can be attributed to several socio-political developments. Following the signing of the Molotov – Ribbentrop Pact in 1939, in June 1940 the three Baltic States – Lithuania, Latvia and Estonia – were occupied by the Soviet Union and incorporated into the USSR as constituent republics. However, in 1941 the Soviet Union was attacked by Nazi Germany, leading to a brief occupation of the Baltic region by the Hitlerite Third Reich. After the retreat of the German armed forces, the Soviets re-established the annexation in 1944. Lithuania would remain a part of the Soviet Union until the adoption of the declaration of independence on March 11, 1990 (Plakans 2011: 343-348).

As a result of these socio-political developments, in the last 70 years Lithuanian Poles have been affected by two powerful language ideologies, Russification (Hogan-Brun and Ramonienė 2003: 29) and Lithuanisation (Chodakiewicz 2012: 424). On the one

hand, cultural Sovietization and linguistic Russification were aimed at assimilating the occupied nations and ensuring the domination of the Russian language, while on the other, Lithuanisation was meant to counteract the previous prevalence of Russian by introducing Lithuanian as the dominant language (ibid). Both of these processes are explained in greater detail below.

1.2.1 Russification

The process of Russification and its impact on the Polish community in Lithuania is addressed in great detail by Kawęcki (2013: 27–28). He specifies that one of the aims of the Soviet authorities was systematic ‘depolonisation’, which was achieved by deliberate demographic shifts in the predominately Polish region of Vilnius by repopulating it with Russians and Lithuanians, as well as substantially limiting the development of Polish-medium education. As Borkowicz explains (1993: 9), the 1960s were characterised by the process of industrialisation, which resulted in thousands of rural inhabitants moving out of the rural areas and settling in the cities. The Vilnius region saw a great influx of ‘highly russified’ Poles from Belarus who replaced the original Polish population. More and more Poles, especially members of the younger generation were starting to use Russian in their everyday lives, whereas Polish increasingly became the language of the older generation (ibid).

Although Polish-medium schools were never banned during the Soviet times, many Poles chose to enrol their offspring into Russian-medium schools. According to Kawęcki (2013: 28), this decision was partly influenced by the popular belief that the level of education was much lower in Polish-medium schools. Russian-medium schools, on the other hand, offered not only higher standards of tuition, but also better career prospects, since fluency in Russian was associated with upward mobility and social advancement.

The fifty-year long period of Russification affected not only the political, social and

cultural life of the Polish community in Lithuania, but also significantly altered its linguistic repertoire. As Hogan-Brun and Ramonienė (2003: 31) explain, following the prolonged period of Russification, “not only people of Russian nationality declared Russian as their first language, but also many members of the ethnic minorities. Russian was also considered to be the first language by a sizeable portion of Poles, Belarusians, Ukrainians, Jews and others who used it as a *lingua franca* in both public and private life [...]. Of those who did not consider the national language¹ to be their first language, some (most Russians) were monolingual; others were bilingual (mother tongue/Russian) or trilingual (mother tongue/Lithuanian/Russian).”

1.2.2 Lithuanisation

Following Lithuania’s independence in 1990, Lithuanian was reinstated as the official state language of the country. While not as numerous as in neighbouring Latvia and Estonia, ethnic minorities were nevertheless regarded as a ‘challenge’ by the new government since they either did not speak the national language or did not master it well enough to use it confidently (Hogan-Brun and Ramonienė 2003: 33). The change in political powers marked a new period in the history of the Polish community, described by Polish historians as ‘the beginning of Lithuanisation’ (Kawęcki 2013; Chodakiewicz 2012). The main aim of Lithuanisation was to counteract the previous prevalence of Russian in public spheres and institutional settings, as well as to ensure that members of ethnic minority groups were able to use the state language with high levels of proficiency. The main target area was the medium of education, hence new curricula, textbooks and other teaching aids, as well as a language-proficiency assessment system for both schools and adult-teaching institutions were developed (Hogan-Brun and Ramonienė 2003: 34).

Interestingly enough, the Polish and Russian minorities in Lithuania are said to

¹By ‘national language’ the authors mean ‘state language’.

display different behaviour towards the idea of integration into the Lithuanian society. While members of the Russian community seem to have accepted the changed socio-political situation and often choose Lithuanian-medium education for their children, members of the Polish community are keen to preserve their cultural and linguistic independence, and are therefore unwilling to integrate with the Lithuanian majority (Juozeliūnienė 1996: 200). A lot of Poles choose Polish-medium schools for their offspring, which may be one of the factors contributing to their lower level of state-language proficiency and academic performance. Hogan-Brun et al. (2005: 354) claim that Polish pupils are for the most part “sufficiently conversant with the state language” to continue their studies in Lithuanian educational establishments; however, their “level of educational achievement actually tends to be lower” (ibid). When describing the socio-demographic situation of the Polish community, Hogan-Brun and Ramonienė note that “[Poles] occupy lower social positions, the greater percentage of them being manual workers” (2003: 37). However, it should be borne in mind that the studies cited here were carried out at least a decade ago, which means that the information that they convey may no longer be entirely accurate. As will be illustrated in the next section, some of the more recent studies suggest that there might have been a change of attitude towards assimilation with the Lithuanian majority, as well as an increase in the state-language proficiency levels.

1.2.3 Linguistic Repertoires

As a result of Russification and Lithuanisation, the younger generation of Lithuanian Poles, i.e. the ones who were born in the Soviet times but educated in independent Lithuania, developed complex linguistic repertoires, consisting of two varieties of Polish (regional and standard), Russian and Lithuanian. While it would be an overstatement to claim that all members of the younger generation of Poles are fully proficient in all four varieties, there is very little empirical evidence available to either

support or refute this observation. It seems reasonable to assume that the regional variety of Polish is considered to be the mother tongue of the great majority of Lithuanian Poles, while Standard Polish is usually learned by those who are educated in Polish-medium schools, though it is not used on a day-to-day basis. Likewise, Russian is normally acquired at an early age through exposure to the Russian media, close contact with members of the Russian diaspora and/or upbringing in bilingual households. Finally, Lithuanian has been taught as the state language since the reinstatement of independence, thus even those members of the Polish community who are educated in Polish-medium schools and who have little contact outside their immediate social network will be proficient in Lithuanian to a greater or lesser degree. Bearing these observations in mind, it is reasonable to assume that each of the four varieties features at least to some extent in the linguistic repertoires of the younger generation of Lithuanian Poles.

While there are few studies that focus specifically on the linguistic situation in the Polish community in Lithuania, some of the existing ones show that the general assumptions presented above do not apply in all instances and that linguistic repertoires are subject to change as a result of different political and social developments. One of the studies that is worth mentioning here is the ‘Multilingual Cities Project’, carried out between 2010 and 2012. Almost 24,000 pupils in primary schools across Lithuania took part in the survey, aimed at establishing their language proficiency levels, language choice, language dominance and language preferences at home (Ramonienė and Extra 2011). One of the findings emerging from this project is particularly relevant for understanding the current linguistic situation in the Polish community, as well as interpreting the data that will be reported in the second part of this study. As Ramonienė and Extra report (2011: 65), although Poles constitute the largest ethnic minority in the country, Polish is the third, and not the second home language reported by pupils of Polish origin. They attribute this to the Russification policy implemented during the Soviet times, whereby many non-Russian ethnic minorities,

including Poles and Belarusians, were fully assimilated into the Russian culture and for many of them Russian became their native language. The authors further report that there is a frequent mismatch between home language and ethnicity, thus Polish ethnicity often co-occurs with Russian and Lithuanian as home languages (2011: 74).

Another study that reports on the linguistic situation in the Polish community in Lithuania is the cross-Baltic survey, carried out by Ehala and Zabrodska (2013). The aim of their study was to evaluate the ethnolinguistic vitality of ethnic groups in Estonia, Latvia and Lithuania. The authors conclude that Lithuanian Poles can be regarded as “the most assimilated linguistically”, for in certain situations they claim to use Lithuanian more often than their native language. Moreover, both Polish and Russian minorities indicated that they aspire to assimilate and expressed a clear desire to belong to the Lithuanian majority (Ehala and Zabrodska 2013: 77–81). The results of this study appear to indicate that over the last decade there has been a shift in attitudes among Lithuanian Poles towards integration, as well as an increase in the use of the state language.

While some of the quantitative studies outlined above offer a good starting point for assessing the linguistic situation in the Polish community in Lithuania, they also carry a certain degree of limitation, stemming from the fact that they are based on self-administered questionnaires and/or qualitative interviews, without any investigation of the spoken language data. As Nortier (2008: 37) explains, the main fallacy of self-administered questionnaires is “the difference between reported and observed linguistic behaviour”, which cannot be verified without access to natural language data. In fact, as will be demonstrated in the data analysis section, the behaviour reported in the surveys is quite different from the linguistic behaviour reflected in the recordings of spontaneous conversations that will be analysed in this thesis. While the Poles may be using Lithuanian more than any other language in some

situations, they keep it strictly separated from their in-group interaction, which is further corroborated by the fact that the presence of Lithuanian in the present corpus is very negligible.

The scarcity of Lithuanian material in the present data is quite surprising. The best explanation as to why this is the case would probably come from the participants themselves if they were asked this question directly during a sociolinguistic interview. Since this type of data could not be collected due to scope and time constraints, the reasons given here are based on general observation of the researcher.

One possible explanation could be the fact that Lithuanian is regarded as the language of a socially more powerful, monolingual majority group, whereas Russian, as well as the regional variety of Polish, are associated with the minority to which the speakers themselves belong. Consequently Lithuanian is employed in more formal situations in which the interlocutors do not know each other, and in certain institutional settings. Russian, on the other hand, is used in informal contexts where the participants are equals, e.g. friends or family members. Moreover, it still acts as the local *lingua franca*, since it is employed when communicating with the Russian diaspora, as well as other ethnic minorities living in Vilnius, e.g. Belarusians and Jews, which contributes to its prestige and favourable status among the respondents.

Since the recordings collected for the present corpus involve members of the same peer group interacting with each other in highly informal settings, one would not expect them to resort to a more formal and powerful variety in order to gain their interlocutor's approval or assert a higher social status. On the contrary, the speakers will attempt to maintain the existing norms of their in-group interactions to ascertain their in-group membership, which means that a strong presence of Russian is to be expected. However, it should be noted that if the recordings were collected in more formal settings, for example between colleagues in the work environment, it is possible that the presence of Lithuanian might have been more significant.

Since the regional variety of Polish is considered to be the mother tongue of the great majority of members of the Polish community and in most instances constitutes the base language for the Polish-Russian language alternation analysed in this thesis, it is important to present an overview of its main features and outline how it differs from Standard Polish.

1.3 The Regional Variety of Polish

The main linguistic variety spoken in the Polish community is a regional dialect of Polish, usually referred to as the Vilnius dialect (in Polish ‘dialekt wileński’), or the northern ‘borderland Polish’. According to Barbara Dwilewicz (quoted in Krzywicki 2005: 94), the contemporary Vilnius dialect developed from the variety of Polish that was once spoken by the intellectuals from the upper classes who moved back to Poland after World War II. Although the dialect is referred to as ‘the Vilnius dialect’, it is not restricted solely to the Vilnius region. The name is normally employed as a generic term to describe the variety of Polish spoken in Lithuania, but it would be a mistake to think that it is uniform across the whole region. Traditionally, one of its particularities, characteristic of many other dialectological situations as well, has been the division between an urban and a rural variety. The urban variety, spoken in Vilnius, was regarded as more ‘correct’ and unified, whereas the rural variety, spoken outside Vilnius, was said to contain a lot of specific phonological, grammatical and lexical forms, which contributed to the common belief that this variety was more ‘incorrect’ (Kurzowa 1993). However, it should be noted that the urban-rural division in the contemporary variety of Polish spoken in Lithuania seems to be less pronounced than reported in the earlier studies, which could be attributed to the rapid urbanisation of the Vilnius region on the one hand, and increased mobility of rural inhabitants towards the capital on the other hand.

The variety of Polish spoken in Vilnius and the Vilnius region has received considerable attention from Polish scholars throughout different stages of its development. Some of the earliest studies were carried out by Nitsch (1925) and Turska (1939/1982) who attempted to document the main features of the language spoken by the Polish community in the Vilnius region. One of the earliest in-depth studies is a collection of articles written by Turska, which were published posthumously in 1982. In her works Turska investigates the features of the Polish language spoken on the territory of some of the former Borderlands, particularly eastern Lithuania and western Belarus, and tries to delimit one dialect area from another, both linguistically and geographically. Another important monograph was published by Kurzowa (1993), who analyses the phonological, grammatical and lexical features of the varieties of Polish spoken in the Vilnius region, as well as the former Polish territories in what is now Belarus. The nature of these earliest works is neatly summarised by Masojć (1996: 25), who explains that these studies are best described as ‘dialectal’, for the researchers were interested in the language of rural inhabitants of the older generation, while the language of the younger generation of speakers, raised in different socio-political circumstances after World War II, was largely ignored.

The more recent works focus on the influence of Russian, Lithuanian and Belarusian languages on the variety of Polish spoken in the city of Vilnius. While the overall aim is still the same, i.e. to compile a list of the most characteristic features of the local variety of Polish and compare it against Standard Polish, there are more and more studies investigating the language of the younger generations of speakers. For example, Geben (2001, 2003, 2008) and Dawlewicz (2000) focus on Lithuanian and Russian influences on the variety of Polish spoken by secondary school and university students in Vilnius. Dawlewicz’s study (2000) aims to analyse Lithuanian borrowings and calques. He compiles a taxonomy of lexical items of Lithuanian origin, as well as classifies them into several semantic groups. However, the major drawback of his study stems from the fact that it is based on data collected via

questionnaires, and thus on the reported, rather than the actual linguistic behaviour of the participants. Similarly, Geben (2001, 2003) analyses Russian and Lithuanian calques in the language of high school students in Polish schools in Vilnius on the basis of data collected via surveys, without any reference to the spoken language data. The methodological premise of all the three studies is essentially the same: respondents are given a lexical item in Standard Polish and are asked to provide a lexical item that they would normally use, or they are given several lexical items referring to the same entity and are asked to select the one that they use in their every day conversations.

An in-depth overview of the Vilnius dialect along with its detailed features would go beyond the scope of this chapter. However, some of the most characteristic phonological and morphological features of the variety of Polish spoken in Vilnius are summarised in Table 1.1 in order to give a very general introduction. The main aim of this brief overview is to distinguish contemporary northern ‘borderland Polish’ from Standard Polish and to list the features that have been reflected in the transcription process, described in greater detail in section 3.3.

	Feature	Example
PHONETIC	/ɛ/ > /ɪ/ or /i/ in unstressed syllables	1. e.g.: [dɛʃˈtʃovɪ] > [dɪʃˈtʃovɪ], [vʲɛˈtʃɔrem] > [vʲɪˈtʃɔrem] 2. in the prefix, e.g.: [ɲɛˈdavɲɔ] > [ɲɪˈdavɲɔ] 3. 3rd p. sg Present and Future Tense, e.g.: [praˈtʂujɛ] > [praˈtʂujɪ], [ˈpujdʒɛ] > [ˈpujdʒɪ]
	/ɔ/ > /u/ in unstressed syllables	e.g.: [rɔˈsɨjski] > [ruˈsɨjski], [tʂɔˈkɔlvʲɛk] > [tʂuˈkɔlvʲɛk]
	/xi/ > /xɪ/	e.g.: [ˈxɪtrɪ] > [ˈxɪtrɪ], [ˈɔʒɛxi] > [ˈɔʒɛxɪ]
	-owi>-u	DAT sg masculine nouns, e.g.: człowiekowi > człowieku
MORPHOLOGICAL	-e/-a/-owie>-i/y	NOM pl nouns (all genders), e.g.: noce > nocy, ręce > ręcy, bracia > braci, królowie>króli
	-ą/-ę>-a	1. ACC sg of feminine nouns and adjectives, e.g.: smutną piosenkę > smutna piosenka 2. 1st p. sg Present and Future Tense, e.g.: mówię > mówia, pójdę > pójda
	-ą>-o/om	1. 3rd p. pl Present and Future Tense, e.g.: oni pracują > pracujo, przyjadą > przyjado/przyjadom 2. INST sg of pronouns, e.g.: ze sobą > ze sobo, ze mną > ze mno
	-ek/∅>-ów	GEN pl (all genders), e.g.: tabletek > tabletków, kin > kinów, procent > procentów
	-em>∅; -am>a	1st p. sg Past Tense, e.g.: słyszałem > słyszał, wróciłam > wróciła
	-emy>-im/ym	1st p. pl Present and Future Tense, e.g.: pójdziemy > pójdzim, bierzemy > bierzym
	-my>-mi	1st p. pl Present Tense, e.g.: używamy > używami
	-ny/-na>-nny/-nna	adjectives ending in -ny, e.g.: drewniany > drewnianny, blaszana > blaszanna

Table 1.1: Selection of important features of regional Polish

1.4 The Regional Variety of Russian

Since the presence of Russian in the present data is quite prominent, it seems reasonable to briefly focus on the variety of Russian spoken in Lithuania. The first point that has to be made is that the available research on this subject is very scarce. The majority of the existing literature focuses on the influence of Russian on the titular languages of the ex-Soviet republics, and not vice versa. The most detailed account of the linguistic variety spoken by members of the Russian diaspora in Lithuania is Avina's monograph (2006), where she analyses the sociolinguistic situation of the Russian community and outlines some of the morphological and lexical-semantic features of the variety that they speak.

The central premise made in Avina's monograph is that the variety of Russian spoken in Lithuania is heavily influenced by Belarusian and Polish, which can be observed on the phonological, morphological, grammatical and lexical levels. It is important to note that Avina does not attempt to state with certainty which of the two languages, Polish or Belarusian, a given feature originate from and discusses both of them together.

As far as phonology is concerned, Avina singles out realisation of the phonemes /t/ and /d/ as soft affricates in words like [tʃetʲə] - *aunt*, [dʒetʲi] - *children*, as well as realisation of the phoneme /çç/ as [tʃ], in such lexical items as [borʲtʃ] for example. She attributes both of these features to the influence from Polish and/or Belarusian.

As far as morphological influences are concerned, the most common tendency is the use of prefixes analogous to the ones found in Polish and/or Belarusian. Quite common are adjective and adverb forms with the prefix 'za-' - *too*, e.g.: 'zatihaja' - *too quiet*, 'zagorjachij' - *too hot* < Belarusian 'zacihi, zagarachy'. According to Avina (2006: 199), these types of derivations are particularly influential due to

the fact that they also perform the same function in Polish and Lithuanian, e.g.: Russian ‘zamnoga’ – *too much*, Polish ‘za duzo’, Lithuanian ‘per daug’ (‘per’ being the Lithuanian equivalent of ‘za-’). Another morphological feature singled out by Avina is the creation of a mixed paradigm when inflecting Russian masculine proper names ending in ‘-a’ or ‘-ja’ (2006: 200–202). This is particularly prominent in the instrumental case, e.g.: ‘s Zhen’kom, Vit’kom, Kolem’. This innovation can be attributed to the influence from Polish, where the respective ending for this group of masculine proper names would be ‘-em’, e.g.: ‘z Witkiem’.

As far as lexical-semantic influences are concerned, Avina (2006: 197) lists lexical borrowings as the most prominent feature. The majority of loanwords have equivalents in Standard Russian and can usually be found in the spoken variety. A lot of these items belong to the categories of household and day-to-day items, e.g.: ‘lahmany’ – *rags* < Polish colloq. ‘łachmany’ or Belarusian ‘lahmany’; ‘fal’bony’ – *frills* < Polish ‘falbany’ or Belarusian ‘fal’bony’; ‘kumpjak’ – *ham* < Belarusian ‘kumpjak’ or regional Polish ‘kumpiak’ (ibid).

Other semantic groups include descriptions of people, usually with negative connotations, e.g.: ‘brudas’ – *a dirty and/or untidy person* < Polish ‘brudas’; ‘gapa’ – *scatterbrain* < Polish ‘gapa’ and family members, e.g.: ‘švager/švagerka’ – *brother-in-law/sister-in-law* < Belarusian ‘švager/švagerka’ or Polish ‘szwagier/szwagierka’; ‘bratova’ – *sister-in-law (brother’s wife)* < Polish ‘bratowa’; ‘testeva’ – *mother-in-law* < Polish ‘teściowa’ or Belarusian ‘cesceva’; ‘matka’ – *mother* < Belarusian colloq. ‘matka’ (Avina 2006: 198).

However, it is important to note that despite these influences, Avina (2006: 79) concludes that the great majority of members of the Russian diaspora communicate using the codified, standard variety. Of course there are speakers who employ the local vernacular as their native language and are not able to use the standard, as well as those who have assimilated with the Lithuanian majority and have partially/mostly

lost their command of the Russian language. Nonetheless, the variety of Russian spoken in Lithuania can be described as very close to the standard variety spoken in Russia.

Chapter 2

Theoretical Background

The aim of the present chapter is to give an overview of the existing literature on the phenomena arising as a result of language contact, which will be discussed in the course of this thesis. Following Auer (1999), bilingual speech will be regarded as situated along a continuum, which typically starts with code-switching, then gradually evolves into code-mixing and finally moves towards a conventionalised fused lect. Since all three phenomena involve alternating between two or more languages, ‘language alternation’ will be employed as a cover term to refer to any or all of these three processes.

The chapter starts with a survey of the existing literature on code-switching, which is one of the most commonly observed and widely researched areas in contact linguistics. The existing literature roughly falls into two main approaches, the grammatical approach and the socio-functional or pragmatic approach. The major findings emerging from both of these approaches will be reviewed in the present section.

Following the direction of Auer’s continuum (1999), the survey will then move towards code-mixing and the related notion of emerging mixed codes. The existing studies report that over the course of time frequent switching between the two languages may become the norm, which means that individual instances of language alternation lose

their local pragmatic force and no longer fit the description of ‘classic’ or prototypical code-switching and should therefore be regarded as an instance of code-mixing. The main aim of this section will be to outline the steps that a given pattern of language alternation undergoes during the transition from switching to mixing, as well as to list the main differences between the two phenomena.

Finally, code-mixing may then develop further into a fused lect, characterised by a loss of variation and grammaticalisation of certain lexical items, such as conjunctions and discourse markers, whereby the entire category from one language is replaced by the equivalent category from another language. This suggests that the three types of bilingual language behaviour are situated along a continuum from pragmatics and discourse to a more grammaticalised system (Walters 2014: 47). The process of ‘fusion’ and the related concept of a fused lect, which are said to be situated at the furthest end of the continuum, will be discussed in the last part of this chapter. This will be followed by an overview of Myers-Scotton’s Matrix Language Frame Model and so-called ‘composite code-switching’. Since Myers-Scotton’s morphosyntactic model was developed as an independent theoretical framework in its own right, i.e. without building upon other existing frameworks, one of the aims of the discussion presented in this chapter will be to explore the relationship between the theoretical concepts proposed by Myers-Scotton and other existing accounts of language alternation research.

2.1 Code-Switching

As the existing research demonstrates, the phenomenon of code-switching (henceforth CS) can be approached from several perspectives. Most of the existing work tends to fall within two major approaches, namely a pragmatic approach, which is concerned with studying the meanings brought about by CS within conversational episodes, and

a grammatical approach, which focuses on analysing the underlying rules, structural regularities and constraints that explain the patterns of language alternation (Gardner-Chloros 2009: 10). While each of these approaches reveals different aspects of CS, one does not exclude, but rather complements the other, for it is the combination of both that provides the more holistic and comprehensive insight. What follows is a more detailed overview of the existing research carried out within each of these approaches, and the major findings that have contributed to our understanding of the phenomenon of CS.

2.1.1 Grammatical Approach

Grammatical approaches have been one of the most prolific sub-fields in the study of CS. According to Gardner-Chloros (2009: 95), we can distinguish between three major trends in the grammatical study of CS, namely variationist approaches, based on universal constraints on where CS can occur in a sentence, generativist approaches, which attempt to explain CS constraints by referring to the Government and Binding theory, and production approaches, concerned with psycholinguistic aspects of CS.

Possibly the most influential study within the variationist approach is Poplack's (1980) research on English-Spanish CS in the Puerto Rican community in New York. It is mainly known for introducing the free morpheme constraint and the equivalence constraint, which are aimed to prove that CS does not occur at random points within a sentence but is subject to certain rules and restrictions. The free morpheme constraint states that a switch cannot occur between a bound morpheme and lexical form unless the latter has been phonologically integrated into the language of the former. Poplack (1980: 234) explains that a form such as '*eat-iendo' - *eating* would be a violation of the free morpheme constraint and would probably not occur in bilingual speech. The Spanish bound morpheme -iendo - *-ing* is affixed to the English root 'eat', which is

a free morpheme. However, according to Poplack, if there is a switch between a free and a bound morpheme, then phonological integration of both morphemes is required. In this case, the free morpheme ‘eat’ would have to be phonologically integrated into the language of the bound morpheme for a switch to take place.

The equivalence constraint states that CS cannot occur at points in the sentence where the surface structures of the two languages differ. Following this simple constraint, a switch is inhibited from occurring within a constituent generated by a rule from one language which is not shared by the other. According to Poplack (*ibid*), a sentence such as ‘El man que came ayer wants John comprar a car nuevo.’ - ‘*The man who came yesterday wants John to buy a new car*’. is only partially correct. It violates the equivalence constraint because it applies an English infinitive complementiser rule, which is not shared by Spanish, to the verb phrase complement. Since a switch did take place in this example invented by Poplack, an English rule was lexicalised in Spanish, resulting in a construction which could not have been generated by a Spanish rule, and which is therefore ungrammatical by Spanish standards. On the other hand, the first portion of the sentence was generated by rules which are shared by English and Spanish, i.e. the L1 and L2 versions map onto each other, constituent-by-constituent and element-by-element. A switch may therefore occur at any point within the main clause and the utterance remains grammatical by both L1 and L2 standards. However, it should be noted that the existing research has provided numerous counter examples where such violations are in fact possible, and thus demonstrated that these constraints do not universally apply to all data-sets (cf. Bhatia and Ritchie 1996; Clyne 1987; Nortier 1990).

Poplack’s work also deserves credit for introducing three types of CS on the basis of the syntactic structures that are involved, namely extra-sentential switching, intra-sentential switching and intersentential switching. As their respective labels imply, extra-sentential switching refers to single lexical items or constructions that remain

outside the syntactic structure of a sentence. They may include tags, discourse markers, agreement particles or short idiomatic expressions. The second type, i.e. intrasentential switching, takes place within a sentence or clause at word boundaries. It can involve single words, as well as several lexical items, as long as they do not arch over the clause or sentence boundary. The third type, intersentential switching, refers to switches that occur at clause or sentence boundaries, where one clause or sentence is in one language and the next clause or sentence is in another language.

In her study of Spanish-English CS Poplack makes a link between the type of switching and language proficiency levels of the speakers who display preference for one type over another. Thus extra-sentential switching is said to require less proficiency in English and can therefore be employed with ease by Spanish-dominant speakers. Intrasentential switches, on the other hand, are said to be employed by only the most fluent bilinguals due to a great syntactic risk involved. This observation is also supported by Lipski (2005), who claims that intrasentential switching is more complex and elaborate than intersentential switching, and can therefore be employed only by the 'true' and 'balanced' bilinguals. While this observation probably may well hold true for a number of language pairs studied in the existing literature, it would be more correct to state that the degree of language proficiency required to perform intrasentential switching depends on the typological characteristics of the languages involved. In such morphologically complex languages as Polish and Finnish, for example, ad-hoc integration of a lexical item from one language into the base in another language would require an advanced degree of proficiency in both languages. On the other hand, in such language pairs as English and Spanish, for example, where there are no grammatical cases involved, a single lexical item from one language can be inserted into a base in another language without any morphosyntactic adjustments, which suggests that it could be achieved by a speaker who is fluent in only one of the participating languages.

Poplack (1980: 589–590) further argues that intrasentential CS should be regarded as a ‘discourse mode’ and that the very fact that the speaker chooses this mode is significant. This observation is also echoed by Myers-Scotton (2005: 117) who believes that in some communities intrasentential switching becomes so common and natural that it can be regarded as an unmarked choice. She makes a further distinction between intrasentential switching and intraclausal switching since the former could be between two clauses in the same sentence. She strongly argues that the best unit of analysis of CS is the clause, and not the sentence. As Myers-Scotton (2005: 240–241) explains, “it is only within the clause that the language varieties involved in codeswitching are in contact [...]. A sentence can qualify as bilingual, but it might be on the basis of having one clause from one language and the other clause from another language. One can say such a sentence shows codeswitching, but it does not show the two languages actually in contact”. She goes on to conclude that “other types of codeswitching may answer various other questions, but intraclausal switching gives us answers to the most interesting questions about the grammar of codeswitching.”

Following Myers-Scotton, the clause will be adopted as the main unit of analysis in the present study. Intraclausal switches will be further subdivided into insertions consisting of single lexical items and multiword units, embedded in the base/matrix language, which include collocations and idiomatic expressions. Collectively, they will be referred to as ‘embedded language islands’. However, unlike Myers-Scotton’s line of reasoning, the analysis presented in the remainder of this thesis does not reject insertions consisting of entire clauses as irrelevant. Chapter 6 will be concerned with switching between clauses, as well as switching between turns.

2.1.2 Socio-Functional Approach

Some of these ‘other questions’ that Myers-Scotton (2005) refers to in the above quotation, and that cannot be answered by focusing solely on intraclausal switching

are addressed within the socio-functional, or pragmatic approach. Gumperz (1982: 131), one of the pioneers in the field of CS research, analyses instances of language alternation in terms of ‘contextualisation cues’, which can be defined as “verbal or nonverbal cues that provide an interpretive framework for the referential content of a message”. They are used by both speakers and listeners “to interpret what the activity is, how semantic content is to be understood and how each sentence relates to what precedes or follows [it]” (ibid). The conversation-related activities singled out by Gumperz include the use of language alternation for quotations or direct speech, addressee specification, reiteration, or message qualification. Most of these activities occur at major syntactic and prosodic boundaries, i.e. at the clause or sentence level. Since switching is activity-related, the utterance units affected by language alternation must be large enough for such an activity to take place, and therefore they normally involve interclausal switches.

Subsequently, Auer (1984, 1995, 2001) further builds on Gumperz’s notion of contextualisation cues within his own Conversation Analytic approach to CS, where he explicitly acknowledges that CS should be analysed as a discourse feature. He puts great emphasis on the importance of sequentiality in uncovering the meaning behind code-switched utterances. According to Auer (1984: 5), a choice of a particular language by one of the speakers exerts an influence on subsequent language choices by the same or other speakers. It is therefore crucial to engage in a detailed turn-by-turn analysis of sequential conversation structures in order to achieve a full understanding of code alternation. He believes that the meaning of CS can best be found at a level of conversational structure, which is neither grammatical nor dependent on larger societal structures, although he acknowledges that both of these can be relevant in the interpretation process.

As Auer and Eastman (2010: 97–98) explain, one of the main premises of the Conversation Analytic approach is that CS occurs in a sociolinguistic context in which

speakers orient towards one language at a time. By departing from the established language of interaction, the speaker signals ‘otherness’ of the upcoming contextual frame and is therefore making a conversational ‘move’, which needs to be interpreted in each individual instance. Some of the possible ‘moves’ may include initial or clarification request, elaboration, indirect decline, setting off direct or reported speech, or repetition.

This sequential approach to CS allows Auer to distinguish between two types of language alternation, namely participant-related and discourse-related switching. Participant-related CS can be a reflection of imbalanced bilingual competence on the part of the speaker, or a manifestation of his preference for one language over another. Discourse-related CS, on the other hand, is triggered by certain changes in the conversational situation, e.g. a change of participant make-up, or an indication of different conversational ‘move’, such as the ones described above (Auer 2001: 444).

Participant and discourse-related switching can be further subdivided along a second variational dimension, namely CS of the ‘alternational’ type and CS of the ‘insertional’ type. As Auer explains (1999: 313), in the alternational type “a return after the switch into the previous language is not predictable”, i.e. there is a change in the base language from the switching point onwards. In the insertional switching, the required lexical items are inserted into a surrounding passage in the other language, while the base language remains the same (Auer 1999: 314). In other words, alternational switching is more ‘permanent’, in the sense that the speakers orient themselves towards one language at a time, whereas insertional switching is characterised by ad-hoc insertions of single lexical items from the other language into the base language.

Auer’s call for a conversation-analytic approach to CS has been employed and developed further by a few other prominent researchers (Sebba 1993; Li Wei 1994, 1998, 2005; Milroy and Li Wei 1995). Li Wei can be credited as one of the main adherents to the

Conversation Analytic approach who has further developed Auer's ideas. He points out that Conversation Analysis "has a specific focus on the interpretive procedures which people routinely draw upon in face-to-face interaction" (2005: 380). Unlike other approaches to the study of CS, it pays little attention to such social variables as gender, age, or occupation, although it does not reject them altogether. According to Li Wei (1998), traditional social variables are 'brought about' in the conversation, rather than 'brought along'. He demonstrates how this can be achieved in his research on CS among English-Chinese bilinguals in the UK. By providing a turn-by-turn analysis of conversational structures and components Li Wei shows how such issues as attitude, preference and community norms are 'brought about', or emerge from the actual contributions of the participants, rather than are 'brought along' into the conversation (1998: 171).

While the Conversation Analytic approach has certainly been one of the most influential approaches to the study of bilingual interaction, it has received its share of criticism. Stroud (1998) criticises conversation-analysts for providing a one-sided explanation of CS by ignoring ethnographically documented information. His main argument is that "conversational code-switching is so heavily implicated in social life that it cannot really be understood apart from an understanding of social phenomena" (1998: 322).

Myers-Scotton and Bolonyai (2001: 4) also criticise the Conversation Analytic approach for its preoccupation with structural features of conversation, which results in "an exceedingly 'flat' explanation of choice". The authors further point out that conversation analysts tend to significantly downgrade the information about conversational interlocutors provided by the wider social context, paying little attention to who participants are in demographic, social-network, and even ethnographic terms (2001: 5). Myers-Scotton and Bolonyai also argue that as a result of their preoccupation with structural elements, conversation analysts ignore the implied knowledge that

speakers develop through their actual conversations. They further specify that this knowledge includes “not just how certain interactions proceed [...], but also the socio-psychological associations and, therefore, the social messages carried by one linguistic choice rather than another” (2001: 5).

Lüdi and Py (2009: 155) point out that one of the weaknesses of the Conversation Analytic approach is its preoccupation with the ‘homoglossic’, monolingual vision, which was possibly influenced by the once dominant idea that an ideal speaker-hearer is someone who lives in a completely homogenous speech community and is therefore a monolingual, first advocated by Chomsky (1965: 3) in the early years of the development of transformational-generative grammar. This critique of the Conversational Analytic approach is also echoed by Auckle and Barnes (2011: 108), who explain that “the theoretical background of constructs such as ‘alternational’ and ‘insertional’ CS is founded on the wisdom that monolingualism usually represents a desired state of proficiency [...]” It should be noted that this critique applies mainly to the earliest accounts of CS, particularly the notion of ‘alternational’ switching, where the speaker switches to a new language of interaction for a longer period of talk and the return to the previous language is not predictable. However, it does not apply to CS of the insertional type, which explicitly endorses the bilingual mode of speaking as the standard.

Despite its minor drawbacks, Auer’s Conversation Analytic account of CS, along with its concepts of participant-related and discourse-related switching, as well as switching of the insertional and alternational type, is often regarded as an example of ‘classic’ CS, or CS of a prototypical kind. Moreover, it serves as a well-established concept against which other phenomena arising as a result of language alternation can be analysed and described. One of these phenomena, which is normally referred to as ‘code-mixing’, or ‘an emerging mixed code’, is discussed in the following section.

2.2 Code-Mixing and Emerging Mixed Codes

The need to account for the varying densities of CS and various structural patterns that it may involve can be regarded as one of the motivations behind different theoretical frameworks that analyse further processes which develop in communities where bilingualism becomes the norm. According to Auer (1999: 314), there are many cases of the juxtaposition of two languages other than CS, one of them simply referred to as ‘code-mixing’ (henceforth CM). There have been numerous examples in the existing research where the observed patterns of language alternation do not fit the prototypical pattern consisting of discourse-related alternational and insertional CS, e.g. Myers-Scotton’s (1993) ‘CS as the unmarked choice’. The existence of this type of language alternation does not mean that CS can only be either marked or unmarked, but rather that it can be driven by a mix of conscious and unconscious choices and that the presence of one does not exclude the other. In unmarked CS, language alternation becomes the usual mode of daily communication, thus individual language choices are not assigned any particular significance. However, in some instances speakers may still employ language alternation with a particular stylistic or pragmatic effect in mind and thus make a conscious decision to switch from one language to another, e.g. to set off direct speech from the main narrative.

According to Auer (1999: 309), different types of language alternation phenomena can be situated along a continuum, which typically starts with prototypical CS, which may gradually develop into CM and then finally morph into a fused lect. In characterising a middle position on the continuum of language alternation phenomena, Auer (1998: 16–21) identifies several steps. The first step involves the abandonment of the preference for one language of interaction. In Auer’s account of CS of a prototypical kind, alternational and insertional types of switching co-exist. Recall from the previous section that in alternational switching a new language of interaction is established from the switching point onwards for a longer period of talk. In

the emerging mixed code, a base language is only interspersed with insertions from another language, though these insertions tend to be rather frequent. Thus according to Auer (1999: 321), the first step is characterised by the abandonment of alternational and insertional types of switching in favour of the insertional type only, which is characterised by the presence of intraclausal switches. However, as Auer himself admits (1999: 315), and as will be demonstrated in the data analysis section of the present thesis, this step is not necessary since the two types of switching can co-exist for a prolonged period of time even in advanced situations of language mixing.

The second step follows directly from the first one. Since in CM switching between varieties becomes very frequent, it starts to lose its contextual power and is seen by the participants as a norm. It is therefore difficult to argue that the juxtaposition of the two languages is triggered by a change in the discourse mode or the conversational situation, hence the second step is characterised by the loss of contextual value and discourse-related functions that can normally be attributed to a given instance of language alternation.

The third step on the continuum from CS to CM is that in some instances it is difficult to establish the language of interaction, which may become ambiguous as a result of frequent language alternation. On the basis of Swiss German and Italian data, Auer (1999: 315) demonstrates that there are instances where it is difficult if not impossible to say whether the language of interaction is Italian or the Swiss German dialect. Instead, it is their alternating use which in itself can be regarded as the preferred discourse mode. Moreover, as a consequence of the frequent intraclausal mixing of the two languages, it is often difficult to maintain the distinction between insertional and alternational juxtapositions. In fact, a typical feature of CM is that alternational and insertional strategies converge. Yet, as Auer (*ibid*) specifies, the distinction does not always collapse completely. It is often possible to identify mixing styles of a more insertional kind and those of a more alternational kind, which further corroborates the

assumption that the transition is gradual and different types of language alternation may co-exist for longer or shorter periods of time.

The fourth and final step may lead to a further development, namely transition from CM to a fused lect, and is linked to the processes of grammaticalisation and lexicalisation, whereby some lexical elements, in particular discourse markers, within the mixed code take on a new grammatical function or a lexical meaning and are no longer regarded as foreign (*ibid*). While the processes of ‘fusion’ and the concept of a fused lect are addressed in greater detail in the following section, it should be noted here that the problem with the scenario proposed by Auer is its preoccupation with the changes in form without paying much attention to meaning. The main diagnostic of a fused lect is based on the study of the processes taking place at the level of grammar, but not much attention is paid to the lexical-semantic properties of the affected items. Since in intense language contact situations both the lexicon and the syntax can be affected, it is important to focus on changes that emerge on both of these levels in the process of transition towards a fused lect.

2.3 Fusion/Fused Lects

An even further development from CM is that of fused lects (Auer 1999) and ‘fusion’ (Matras 2000). According to Auer (1999: 321–323), following a longer period of heavy mixing between several languages the use of one language or the other for certain constituents becomes obligatory. The transition from CM to a fused lect starts at what Auer calls ‘relatively unbound elements of grammar’ (1999: 324), which include discourse markers, conjunctions, and certain adverbials. These lexical elements typically serve the function of modifying the utterance, and as such, they perform a discursive function rather than a referential one. These unbound elements are usually affected by structural sedimentation sooner and faster than other elements,

which suggests that the transition from language mixing to fused lects for the unbound elements does not tend to coincide with the transition for the other structural domains (ibid).

Auer's observation regarding the unbound elements of grammar and their role in the transition from CM to a fused lect originates from Matras' account of 'fusion'. According to Matras (1998, 2000), in communities where bilingualism has become the norm, speakers may over time adopt just one set of particular linguistic structures, thus entirely eliminating or partially reducing the need to select among competing word-forms. This is usually the case when discourse marking items are involved, which in Matras' account include conjunctions, tags, interjections, and focus particles (1998: 144–145). According to Matras (2000: 516) discourse markers are elements that help speakers monitor and direct the way a propositional content of the utterance is processed and accepted by the hearer. The monitoring and directing operations are cognitively more complex than formulating propositional content, for the latter involve back-processing, planning ahead, anticipating and controlling reactions, interpreting gestures and intervening with a possible undesired course of processing on the part of the speaker (Matras 2000: 517). Therefore, bilingual speakers will aim to reduce this cognitive load by showing preference for just one set of discourse-marking devices and abandoning the other.

A point should be made about the term 'fusion' itself. Since lexical items affected by the process of 'fusion' are extrasentential, it could be argued that this is not a sufficient basis for calling the whole system 'fused'. While in Matras' account 'fusion' revolves solely around discourse marking lexical items, Auer's approach is a bit more inclusive. He acknowledges that in addition to the unbound elements of grammar, content words may also be affected in the process of transition from CM to a fused lect. This process is known as 'borrowing' and involves integration of lexical elements into the receiving variety, leading to the gradual exclusion of the equivalent forms in the

surrounding language. However, as he further specifies (1999: 327), “such borrowing alone is not enough to speak of a fused lect since it only superficially affects the language system (i.e. its lexicon) but as such has no consequences on the grammar”. It is therefore crucial to look for other signs of convergence which affect not only the lexicon, but also feature on the grammatical level in order to label a given system as an instance of a fused lect.

2.4 Matrix Language Frame Model and Composite Code-Switching

The processes of borrowing, convergence and language alternation are also at the core of Myers-Scotton’s account of so-called ‘composite CS’, which is reminiscent of what other researchers refer to as CM, as well as to some extent Auer’s fused lects. As part of her Matrix Language Frame Model, which will be explained in more detail in Chapter 5, Myers-Scotton (2002: 8) makes a distinction between ‘classic CS’ and ‘composite CS’. Following this theoretical framework, ‘classic’ CS is characterised by two languages in contact with each other. However, only one of them, labelled the Matrix Language, supplies the morpho-syntactic frame for the bilingual speech, whereas the other one, labelled the Embedded Language, provides some of the lexical content. Moreover, ‘classic’ CS is concerned with only one type of switching, namely intraclausal switching, since it is the only type of language alternation where the grammatical systems of the two languages are actually in contact with each other (Myers-Scotton 2005: 241).

Two remarks have to be made with regard to Myers-Scotton’s definition of CS. The first one is that it attempts to investigate CS from a purely grammatical perspective, hence explicitly rejecting any instances of language alternation that go beyond the clause level. The second remark, which follows directly from the first one, is that

her definition of ‘classic’ or prototypical CS is qualitatively different from the one outlined in the earliest studies of language alternation (cf. Auer 1984). Recall from the previous section that in the Conversation Analysis approach CS of the prototypical kind is made up of both insertional switching of single lexical items and alternational switching of larger constituents, such as entire clauses or sentences, which may also perform a discourse-related function. Subsequently, an increase in frequency of intraclausal switches may lead to the transition into CM. This suggests that the phenomenon that Myers-Scotton refers to as ‘classic CS’ partly overlaps with what other researches would classify as CM, as long as the inserted items follow the morpho-syntax of the matrix language. It should be borne in mind that the definition of ‘classic’ or prototypical CS which will be used in the present study is that of Conversation Analysis, i.e. ‘classic CS’ will be approached from a more holistic perspective, taking into account its grammatical properties, as well as analysing its discourse-related or pragmatic functions. In other words, the notion of ‘classic CS’ adopted in this thesis does also, unlike Myers-Scotton, accept interclausal switches to be relevant phenomena of CS.

The other concept proposed by Myers-Scotton to be discussed here is that of ‘composite CS’ (2002: 8). Just like ‘classic’ CS, this type of language alternation also contains surface level morphemes from two or more languages. However, Myers-Scotton (2002: 242) specifies that “composite CS also shows convergence in regard to the source of some frame-building procedures, as well as in the features of the abstract grammatical structure in some lexemes.” She further suggests that composite CS should be regarded as a combination of CS and convergence. The latter can be defined as “a linguistic configuration with all surface morphemes from one language, but with only part of its abstract lexical structure from that language, and the rest from another. [...] this means that the abstract grammatical structure of the receiving language (its lexical-conceptual structure, its predicate-argument structure, or its morphological realisation patterns) becomes a composite, even though its surface appears uniform”

(Myers-Scotton 2003: 85). Thus similarly to ‘classic’ CS, in composite CS one of the participating languages contributes most of the morpho-syntactic structure. However, the main difference between the two types lies in the fact that the other participating language also contributes some of the abstract structure underlying surface forms in the clause. As Myers-Scotton explains (2005: 240–241), the abstract structure refers to all the grammatical rules that need to be met for a clause to be well-formed. On the other hand, surface structure refers to the actual linear order of words. What this means for language alternation is that in composite CS not only surface-level morphemes originate from both languages, but also the abstract structure of the frame itself. In this respect, composite CS is closely related to what other researchers define to as CM, or an emerging mixed code.

The Matrix Language Frame Model, as well as the critique that it has received are addressed in greater detail in section 5.2 in Chapter 5, it should be noted here that the model is particularly appealing since it aims to address the apparently very heterogeneous phenomena emerging from language contact within a single theoretical framework. While Auer specifies that one of the characteristics of an emerging mixed code is that the base language may be ambiguous, he does not state explicitly how this can be assessed. Myers-Scotton’s model, on the other hand, can be particularly useful in this respect since it offers a theoretical framework for evaluating the level of structural convergence between the two languages in contact and its possible manifestations on the level of the morpho-syntactic frame.

Following the MLF model, the study of language alternation presented in this thesis is concerned not only with lexical CS on the surface level, but also other contact-induced phenomena, which result in a change in the structure of the receiving language, namely lexical and semantic borrowing, loan translations and larger structural calques. The theoretical position taken in the present data analysis is based on the idea that both overt lexical and covert syntactic phenomena should be analysed together under

one unified framework in order to give a through account of the outcomes of prolonged language contact.

While the contact-induced phenomena mentioned above can all be classified as instances of either lexical or structural borrowing, it should be mentioned that there are also changes that arise as a result of language contact, but that do not relate directly to the structure of the other language. Thomason (2006: 342) describes them as “contact-induced changes that do not involve direct interference” since they do not result in the diffusion of any lexical or structural features from one language to another. As a result of such processes as attrition, language death or imperfect acquisition due to the lack of input, speakers may over time simplify case systems or develop limited registers in a given language. Since these outcomes arise as a result of language contact and can therefore be classified as contact-induced change, it would be reasonable to include them under the same framework as CS and various types of borrowing.

While there is no evidence to suggest that the regional Polish case system has started to erode as a result of the influence from Russian, the same cannot be stated with certainty with regard to speakers’ preferences to employ certain syntactic structures. One such example is the speakers’ preference to use the more analytic construction ‘dla’ followed by the genitive case, instead of the dative case, which is normally preferred in Standard Polish, e.g. ‘dla Jacka’ instead of ‘Jackowi’ - *for Jacek*. However, the problem with this example is that it is not clear whether the preference for this particular construction originated as a result of the influence from Russian. Both of these constructions also feature in Standard Russian and there is no evidence to suggest that one of them is preferred over the other. Moreover, the preference for ‘dla’ followed by the genitive case has also been reported for other dialects spoken in eastern regions of Poland.

The observations presented above suggest that the presence of contact-induced changes

that do not involve direct interference should not be dismissed too soon. However, they can be quite difficult to identify and assess properly on the basis of a very limited corpus of data. Moreover, the origin of any potential examples would have to be verified with a more thorough investigation, which would go beyond the scope of the present study. For these reasons they will not be addressed in greater detail in this thesis.

Chapter 3

Research Questions, Data and Methodology

One of the aims of the present chapter is to outline the hypotheses which will be tested in the thesis. While the main focus is on an emerging Polish-Russian mixed code, the extent to which ‘classic’ or prototypical CS features in the corpus will also be analysed, for the presence of one does not exclude the other. Moreover, this chapter will attempt to define two concepts that are at the centre of the methodological framework, employed in the present study. The first one is the concept of the speech community, while the second one is the concept of the social network. This will be followed with a description of the process of data collection and selection of informants. Finally, the rationale behind the transcription technique used in the present study will be explained.

3.1 Hypotheses

As stated in the introduction, one of the main aims of the present study is to establish to what extent the practice of regular language alternation in the Polish community

in Lithuania can be classified as an instance of an emerging mixed code, and to what extent it can be regarded as an instance of ‘classic’ CS. According to the existing literature, the presence of one does not exclude the other, for the two phenomena might co-occur for a longer or shorter period of time. The principle aim is therefore not to make any categorical claims but rather to establish which types of language alternation phenomena can be identified in the present data and to assess which one of them appears to prevail.

The set up of the chapters for the present thesis is based on Poplack’s (1980) typology of different types of switches, namely extra-sentential, intrasentential and intersentential CS. As their respective labels imply, extra-sentential switching refers to single lexical items or constructions that remain outside the syntactic structure of a sentence, such as tags, discourse markers, agreement particles or short idiomatic expressions. The second type, i.e. intrasentential switching, takes place within a sentence or clause at word boundaries and can involve single words, as well as several lexical items, as long as they do not arch over the clause or sentence boundary. The third type, intersentential switching, refers to switches that occur at clause or sentence boundaries, where one clause or sentence is in one language and the next clause or sentence is in another language. This typology is widely accepted within the existing literature on CS and offers an opportunity to start with grammatically more simple to grammatically more complex and relatively more advanced types of switches.

An alternative, and perhaps more exploratory approach would be to structure the thesis on the basis of different types of language alternation and thus start with analysing instances of ‘classic’ CS, then proceed to CM and finally look into evidence for the emergence of a fused lect. In hindsight, the latter approach would have probably worked better than the former since there is evidence in support of all three types of language alternation. However, in the initial stages of the present study the analysis was mainly focused on the grammatical aspects of what seemed to be ‘classic’

CS, hence the original set up worked quite well. Only later did the complexity of the data, especially on the structural level, started to emerge and the presence of different types of language alternation began to come forward. The optimal solution to this issue was to tie each type of switching to a respective type of language alternation, thus extra-sentential and frequent intrasentential switches are said to be characteristic of CM and fused lects, whereas intersentential switches are normally associated with ‘classic’ CS. Each type of switching is addressed in a separate chapter.

As some of the existing literature suggests, most notably Auer 1999, emerging mixed codes are characterised by regular and frequent intraclausal switches of the insertional type from one language into the base of another language. Moreover, as indicated by Myers-Scotton (2002), they are also accompanied by the process of convergence, where the influence of the embedded language manifests itself not only on the surface level, but also in the underlying abstract structure. Another factor that indicates the presence of a mixed code and its possible transition towards a fused lect is ‘fusion’ at the level of discourse markers. In communities where bilingualism becomes the norm discourse marking operators from both languages may be employed at the same time. However, over time, one set of discourse marking operators may be abandoned in favour of the other, which would suggest that lexical items, which were initially imported as switches, have taken on new grammatical functions within the mixed code.

Drawing on the observations outlined above, the following hypotheses will be tested in the present study:

1. The Emerging Mixed Code/Fusion Hypothesis

the pattern of language alternation, observed in the present data can be classified as an instance of an emerging Polish-Russian mixed code if:

- (a) Russian discourse markers can appear as single switches in the Polish base

without triggering further switching. If Russian discourse markers have no competing equivalents from Polish and are exclusively employed in both Polish and Russian contexts, this can be taken as one piece of evidence for the emergence of a fused lect;

- (b) intraclausal switches, consisting of single lexical items and larger Russian constituents, which will be referred to as ‘embedded language islands’, constitute the preferred type of language alternation and are regularly and frequently inserted into the Polish base;
- (c) there is evidence of convergence of the regional variety of Polish towards Russian on the grammatical level, which manifests itself in the presence of structural and single-word calques.

2. ‘Classic’ CS Hypothesis

the pattern of language alternation, observed in the present data can be classified as an instance of ‘classic’ CS if:

- (a) there is a clear preference for one language of interaction;
- (b) there is a change in the base language from the switching point onwards, and the return towards the previous language of interaction is not predictable;
- (c) it is possible to identify at least a tentative discourse-related function that CS performs;
- (d) there is no evidence of convergence of the regional variety of Polish towards Russian on the grammatical level, which manifests itself in the presence of structural and single-word calques.

The first hypothesis will be addressed in Chapters 4 and 5. The analysis presented in Chapter 4 will involve seven lexical items that are most frequently employed

to perform discourse marking functions in the present data, while Chapter 5 will deal specifically with intraclausal switches of the insertional type and will take into account both single switches and embedded language islands, i.e. Russian phrases and collocations, framed by the Polish clause. Single word and structural calques will also be analysed in the same chapter in order to assess the evidence for the presence of convergence. The second hypothesis, which relates to ‘classic’ CS will be tested in Chapter 6, which deals specifically with switches between larger constituents, namely two main clauses, or a main clause and a subordinate clause both within a turn and between turns.

In order to test the hypotheses outlined above, the analysis will focus on the recordings of spontaneous conversations, featuring members of the Polish speech community in Vilnius. However, before proceeding to describe the process of data collection and transcription, it is important to clarify several concepts, namely that of the ‘speech community’ and ‘social network’.

3.2 Methodological Framework

3.2.1 Speech Community

Following Hymes’ position that “the natural unit of sociolinguistic taxonomy (and description) is not the language but the speech community” (1972, quoted in Wolf 2001: 15), it is crucial to delimit and define the groups of speakers within the Polish minority in Lithuania that will be analysed in the present study.

Although the notion of ‘speech community’ is widely used in sociolinguistic research, there is an ongoing debate regarding the definition of this term. As Hudson (1996: 27) points out, although the definitions range from very simple to very complex, there is no need to try to reconcile them with one another as they are simply trying to reflect

different phenomena. He further adds that all of the definitions can be regarded as ‘correct’, since each of them allows us to define a set of people who have something in common linguistically (ibid).

One of the simplest definitions is that of Lyons (1970: 326), who claims that a speech community is “all the people who use a given language (or dialect)”. This definition presumes that speech communities may overlap if some of the speakers are bilingual, and it does not take into consideration such factors as social or cultural unity. A more complex definition is proposed by Hymes (1974: 51), who defines a speech community as “a local unit, characterized for its members by common locality and primary interaction”. The main emphasis in Hymes’ definition thus falls on the presence of a territorial basis and a continuous interaction between members. Hansen (1987: 16, quoted in Wolf 2001) defines a speech community as “a group of people who regard themselves as using the same language”. He further specifies that the speakers are conscious of using the same language, and it is this consciousness that creates a sense of unity within the community. Likewise, Milroy and Gordon (2008: 135) believe that “speech communities [...] are characterized by mobility and access to knowledge of the linguistic and social practices of others”. They further specify that in order to describe a speech community an analyst must define a particular geographical location and then distinguish a series of possibly relevant social categories, e.g. gender, class, or generation cohort, since these categories might influence the behaviour of speakers (Milroy and Gordon 2008: 134). Finally, it is worth mentioning one more important feature of speech communities that was pointed out by Bolinger. According to Bolinger (1975: 333), “there is no limit to the number and variety of speech communities that are to be found in society”. What this means is that any population may contain a very large number of speech communities, with overlapping memberships and overlapping language systems (Hudson 1996: 26).

What follows from these definitions is that a speech community can be defined as

a group of people who are using the same language/dialect (or languages) and are conscious of it. They are bound by a common locality as well as a shared knowledge of the linguistic and social practices, which creates a sense of unity within the community. There is no limit to the number of speech communities since community memberships might overlap.

Bearing these theoretical aspects in mind, the Polish population living in Lithuania can be divided into several speech communities. In the most general sense, all the Lithuanian Poles can be said to form one big speech community. They have a strong territorial basis, which can be specifically described as the city of Vilnius, and the villages of the Vilnius and Šalčininkai regions. They all speak a common variety, i.e. the regional variety of Polish, and they interact with other members of the speech community on a regular basis. This interaction can take place in various social settings, such as the family, school or other educational establishments, sometimes workplace, local neighbourhood, church, existing social groups etc.

However, if we were to narrow down the geographical location, the Polish community in Lithuania can be further divided into two other speech communities, namely an urban Polish community and a rural Polish community, though it should be noted that this separation is increasingly becoming more symbolic, rather than factual. Nevertheless, it will be maintained in the present study for several reasons. The first reason is that the majority of the existing literature, investigating the variety of Polish spoken in Lithuania, dates to the times when the division was still very much relevant. Since the present study aims to further build on the existing research, it makes sense to indicate precisely which part of the Polish community it is concerned with by referring to the well-established terminology. The second reason is that this division can be helpful in defining the speech community that will be investigated in the present thesis. There are still speakers of what is traditionally described as ‘the rural variety of Polish’ living in the villages of the Vilnius region, albeit in very small

numbers, hence it is important to state explicitly that they are excluded from the present study.

Although the division between a rural and an urban Polish community will be maintained in the present study, it should be acknowledged that there is a lot of interaction and mobility between the two groups. While following the collapse of the Soviet Union there was a lot of migration from rural areas to Vilnius due to low and insecure income from farming, in the last few years more and more Poles chose to move out of the city and settle in bigger houses in the rural areas. This was largely facilitated by the improved transport infrastructure, making the daily commute to Vilnius for work purposes possible. Following the traditional division, in categorical terms the urban Polish community would include all the Poles centred in the city of Vilnius. Although there are no areas within the city that are exclusively inhabited by the Poles, the majority of Poles live in the suburbs, as opposed to the city centre. Some of the districts with a large concentration of Poles include Pašilaičiai, Šeškinė, Fabijoniškės, Justiniškės, Viršuliškės, Lazdynai, and Naujininkai. All of these districts are the so-called ‘new districts’ that were built in the 1970s by the Soviets for the working classes, including newly-arrived migrants from rural areas.

In the context of the Polish minority the social categories of age and educational background play an important role. On the basis of these categories the Polish community can be divided into two groups: the older generation, namely those who grew up and were educated in Soviet times (1940 – 1990), and the younger generation, i.e. those who grew up and were educated in independent Lithuania (from 1990 onwards). This division is important because the linguistic repertoires of Lithuanian Poles largely depend on what generation one belongs to and what education one received.

The Soviet education system was aimed at encouraging the use of Russian and suppressing the use of native languages (in most cases Lithuanian and Polish), while

the Lithuanian education system puts an emphasis on the Lithuanian language and culture. Thus, on the basis of age and educational background, as well as on geographic and linguistic grounds the Polish speech communities can be further divided into the younger and older rural Poles, as well as the younger and older urban Poles. The label 'young' is a reference to the post-Soviet generation, who are now not older than their early thirties, and the label 'old' is employed to describe those who were educated in the times of the Soviet Union, who are now in their late thirties – early forties or older.

The analysis of language alternation among Lithuanian Poles, presented in this study, focuses on the younger generation of urban Poles, who are now in their twenties. This speech community makes an interesting research case because it can be regarded as caught between two ideologies: past Russification and current Lithuanisation. Although young Lithuanian Poles grew up and were educated in independent Lithuania, and therefore encouraged to develop fluency in the state language, it can be argued that they were nevertheless influenced by the Russification policies. The long-term effects of cultural and linguistic Russification that lasted for 50 years are difficult to revoke, and although the younger generation of Lithuanian Poles might not have been subject to its influences directly, their parents and grandparents certainly were. A lot of them were educated in Russian-medium schools, which means that they are more proficient in Russian than Polish and would therefore code-mix when speaking with their offspring. Therefore, Russian is normally acquired at an early age through upbringing in bilingual households, exposure to the Russian media, which seem to be extremely popular among members of this age-group, as well as close contact with members of the Russian diaspora. On the other hand, the majority of Poles consider Polish to be their native language and employ it alongside Russian and Lithuanian on a daily basis.

3.2.2 Social Network

While the notion of the speech community was employed to define and delimit the group of speakers chosen as the informants, the concept of a social network was used as a tool for recruiting and identifying potential participants who would be willing to take part in the present study.

A social network is often defined as a web-like pattern of relationships among individuals (Gelles and Levin 1999, quoted in Daming et al. 2008: 262). Two broad approaches to the description and analysis of these relationship patterns can be identified: a sociological approach focusing on the whole network, and an anthropological approach focusing on the egocentric (personal) network (ibid).

In sociolinguistic research the concept of social network is employed both as a method of data collection, and as a theoretical framework for data analysis. As Milroy and Gordon point out (2008: 124), language networks have been successfully employed in investigating the social mechanisms of language maintenance and shift in bilingual communities. They further explain that when language ties remain strong, minority languages are able to resist the institutional pressure, but once they become weak, language shift is likely to take place.

One of the studies employing the concept of social network to analyse language choice and language shift is Li Wei's (1994) account of the Chinese community in Tyneside. Li Wei focuses on three-generation Chinese families and investigates their social networks in terms of ethnic and peer group content in three different types of contacts: exchange networks, interactive networks and passive networks. Exchange networks include family members and close friends, with whom an individual interacts routinely and exchanges direct aid, advice, support etc. Interactive networks are characterised by frequent contact over a long period of time, yet one does not rely on them for personal favours and other material resources, e.g. a tie between a shop

owner and a customer. A 'passive' type of network tie mainly includes physically distant relatives or friends, who offer influence and moral support in spite of absence of regular face-to-face contact.

For each type of network, Li Wei has calculated an ethnic and peer index, i.e. the number of ethnic and peer ties. He found that the ethnic index was the best predictor for an individual's language choice: the more Chinese contacts a speaker has, the more likely he/she is to speak Chinese all the time. Peer index, on the other hand, worked in very complex ways. The more peer-group contacts, i.e. members of the same generation, a grandparent had, the more Chinese was used, and the more contacts a parent had with members of the grandparent generation, the more Chinese he or she used. On the other hand, the more contacts a parent had with members of the child generation, the more bilingual he or she was, while the more peer-group contacts a child had, the more English he or she used. Thus peer content contributed to language maintenance within generations, and to language shift across generations (Li Wei 2008: 270).

The principle aim of social networks in the present research is to use it as a tool of getting access to the speakers and the data. Once several individuals from the relevant speech community were identified, following the 'snowball sampling' technique (Scott 2000; Lanza and Svendsen 2007) they were asked to nominate other participants whom they have regular contact with and who would be willing to participate in the study. As opposed to Li Wei's focus on three types of networks (exchange, interactive and passive), the aim of the present study is to focus on peer ties within the exchange networks, which will help to investigate the patterns of language alternation within a clearly defined unit of study and also focus on variation between individual language practices.

This method of recruitment of research participants proved to be successful, since it allowed to gather a representative sample of speakers who are in regular contact

with each within the target population of interest. One of the concerns related to employing the ‘snowball sampling’ technique is whether a qualitative study based on a small sample size can be used to apply findings to the more general population. However, same as many other qualitative studies of the similar kind, the present investigation does not aim to generalise to the entire population, but gaining a thorough understanding of the sample population, which can be described as ‘the younger generation of urban Poles’. In this respect the use of the ‘snowball sampling’ technique has the considerable advantage since it allowed the researcher to obtain a coherent sample population from among the entire community of Polish speakers. This is particularly important when investigating the phenomenon of language alternation, as CS and CM can produce highly variant practices.

The advantage of using social networks as a methodological tool is that it allows the researcher to identify pre-existing groups of speakers with an established history of prolonged communication, which suggests that linguistic behaviour of the participants is more natural and spontaneous. Moreover, by using the group dynamics the researcher can obtain larger amounts of spontaneous speech. Secondly, it offers a possibility of dealing with variation between individual speakers, rather than between groups constructed with reference to such predetermined social categories as class, ethnicity or gender, which allows to discover the social categories and group affiliations that participants themselves find meaningful.

3.3 Data Collection

The main source of data for the present study are recordings of naturally-occurring spontaneous conversations that feature members of the Polish community in Lithuania in their twenties. Following the guidelines set out by the University of Oxford’s Central University Research Ethics Committee (CUREC), all participants were given

a detailed description of the study (Appendix A.1), outlining the research questions, the format of the recording, information about data protection, and contact details. They were also given a consent form to sign (Appendix A.2), which also explained that they were free to withdraw at any point. In addition, a questionnaire (Appendix A.3) was administered in order to obtain some background information about the informants. The questionnaires were mainly aimed at achieving consistency across the sociolinguistic profiles of the respondents, but did not feature directly in the data analysis.

3.3.1 Informants

Informants for the present study were selected on the basis of three criteria, namely ethnicity, age and geographical location. The criteria were selected in order to delimit a smaller part within the Polish speech community in Lithuania as a whole and to investigate its language alternation practices in greater detail. In order to fulfil the first criterion, the informants were asked to specify their ethnic background in the initial questionnaire. This was done to ensure that there were no respondents who identified themselves with any other ethnic groups that often interact with the Polish community in Lithuania, in particular Belarusians. The second criterion was age, which was delimited to 20–30 years old. One of the reasons for targeting this particular age group is that these speakers are most susceptible to switching due to their direct and indirect experience of the Soviet language policies, partly through the experience of their parents and older siblings, partly through their early education and contact with Russian peers, e.g. in mixed schools or kindergartens, followed by the implementation of post-Soviet policies of Lithuanisation. The third criteria is concerned with location, and aims to ensure that all the informants originate from and reside in the city of Vilnius, as opposed to the rural areas of the Vilnius region, which also have a high proportion of the Polish population.

Informants were recruited via two routes. The first route involved establishing initial contact with several members of the target group during the initial stage of data collection, which took place in September 2011. Following the ‘snowball sampling’ technique, these informants were asked to nominate other members of their social network whom they have regular contact with and who would be willing to participate in the recordings of spontaneous conversations. A total of eleven speakers, four female and seven male, were recruited during the first stage and six recordings were obtained, with one speaker taking part in two recordings made on two separate occasions with two different interlocutors. Participants were left in charge of the recorder and recorded themselves at the time that was most suitable to them.

The second stage of data collection took place in September 2012. This time informants were recruited with the help of social media. A message was circulated on the Facebook group of the Polish community in Lithuania called ‘Pulaki z Wilni’ (*Poles from Vilnius*) which has almost 3,000 followers. Members who expressed their interest in taking part in recordings were contacted, and, once they and their nominated interlocutors answered the questionnaires and gave their consent, they were given a digital recorder so that they could make a recording. The second stage of data collection involved sixteen informants, six female and ten male, and resulted in another nine recordings. The total number of informants who participated in the present study amounts to 27 speakers. However, one recording was dismissed due to its very poor quality, hence the final corpus consists of fifteen recordings and the number of informants includes 25 speakers, ten female and fifteen male. There are two recordings which feature three speakers, while all the remaining recordings feature two speakers. All of the respondents confirmed that they are from Vilnius and identified themselves as ‘Polish’. The average age of the speakers was 24 years old, with the oldest speaker aged 27 and the youngest speaker aged 21.

3.3.2 Recordings of Natural Language

The purpose of the recordings was to obtain natural, spontaneous spoken language data; therefore, informants were left in charge of the recorder and were asked to record a conversation of up to 90 minutes long at the time that was convenient for both of them. The majority of recordings were made with a Sony® ICDPX333.CE7 PX Series Digital Voice mp3 recorder. The resulting recordings were of varying quality, which was largely dependent on the place where they were made. Conversations recorded at home are of the highest quality, with no background noise. On the other hand, conversations recorded in public places, such as cafes and bars feature a lot of background noise, although the speakers' words are nevertheless clearly audible. Two speakers made their recording using the software installed on their computer at home but this did not have any impact on its quality. The data were then exported onto a computer and converted into waveform audio format (.wav) for easy accessibility. With one file per recording session, this resulted in fifteen .wav files, totalling over fourteen hours (14:13:23) of recording time and producing a corpus of 109,628 orthographic words.

3.4 Transcription Technique

When it comes to transcription, an infinite range of styles and degrees of phonetic accuracy are possible. For the purpose of the present study, transcription is based on the conventions introduced by Conversation Analysts, particularly that of Atkinson and Heritage (1984) (see page xvi for a full list of transcription symbols employed in this thesis). This method was first adapted for studies of CS by Auer (1984) and was later followed by other researchers (cf. Li Wei 1994; Sebba 1993). Following this methodological approach, transcription involves writing down as closely as possible not only the actual words uttered by speakers, but also such additional information as

precise beginning and end-points of turns, duration of pauses, overlapping segments of speech, as well as audible sounds other than words (breathiness, laughter) (Hutchby and Wooffitt 2008: 71). In addition, depending on the overall research purpose, some researchers choose to include prosodic features, namely word stresses and intonation patterns (cf. Maschler 2009). Due to time constraints, prosodic features have not been represented in the present corpus. However, it does not mean that they were not taken into consideration. As will be demonstrated in Chapter 5, in some examples word stress can serve as an additional factor in establishing whether a switch from Russian has been integrated into the Polish base. When this is the case, it is stated explicitly if the given lexical item follows Russian stress patterns or if it is stressed on the penultimate syllable in accordance with Polish prosody.

It should also be noted that the labels ‘Polish’ and ‘Russian’ are employed here as a shortcut term to refer to the variety of Polish and the variety of Russian spoken in Lithuania. For the transcription of the Russian spoken data, transliterated Latin script is employed in order to make the examples more accessible to the readers who are not familiar with the Cyrillic script. For the transcription of Polish and Lithuanian data standard orthography is employed. However, in order to set off the variety of Polish spoken in Lithuania from Standard Polish some of the most characteristic features of the regional variety have been reflected in the orthographic transcription (cf. Table 1.1 for a full list).

The regional Polish features selected for transcription, such as the use of non-standard endings in some forms of the verb, (e.g.: first person past tense ‘słyszał – słyszała’ – *I heard* (masculine-feminine), instead of the standard ‘słyszałem – słyszałam’), or the plural genitive (e.g.: ‘zabawków’ – (*of*) *toys* instead of ‘zabawek’) are not only phonetically prominent and therefore easy to identify, but are also consistently produced by all the speakers in the corpus. Other features, such as labialisation of the phoneme /o/, are not only subject to variability, but are also problematic to pick

out in spontaneous spoken speech. Therefore it seemed plausible to reflect only those features that can be unambiguously identified and that feature in the discourse of all the respondents. Moreover, the possibility that some speakers will resort to Standard Polish is kept open, and should it prove to be the case, the distinction between the two will be apparent only if the regional features are reflected in the orthographic transcription.

Since Polish and Russian are two closely related languages, some lexical items are identical in both languages. They include pronouns, e.g. ‘ja’ – *I*, ‘on’ – *he*, verbs ‘był’ – *was*, nouns ‘dom’ – *house* or numerals, e.g. ‘dwa’ – *two*. With regard to the present data, further ambiguity is caused by two additional factors. The first one is the fact that the pronunciation of the regional variety of Polish spoken in Lithuanian has over time diverged from Standard Polish and acquired some phonetic features characteristic of East Slavic languages. The second one is the fact that we are dealing with spontaneous spoken language data, which exhibit reduction from the canonical, careful pronunciation in the form of segment or syllable deletion, significant change in vowel quality or realisation of expected stops as approximants (Warner 2012: 623). As a result of these two factors, there are numerous instances in the corpus where it is problematic to decide whether a lexical item should be transcribed as Polish or Russian. Pronunciation and stress helped to resolve the ambiguity issue in some cases, e.g. with the pronoun ‘ona’ – *she*, since in Polish it is articulated as [ˈɔna], with the stress on the initial syllable, whereas in Russian it is pronounced as [ɐˈna], with the stress on the last syllable. However, as mentioned in the introduction, the criterion based on the analysis of phonetic and prosodic features cannot be applied in every instance and some lexical items will inevitably remain ambiguous. Overall, there are 637 such instances across the corpus. When this is the case, both variants are given, e.g. ‘Mówili że on/ [RUS] on swoju mašinu prodal’ - *[They] said that he sold his car*.

To ensure that examples are easily accessible to all readers, a basic word-for-word gloss and a free translation have been provided below. Language alternation has been picked out by the use of bold script for both Lithuanian and Russian switches. In order to differentiate between the languages involved, switching involving Russian was signified by [RUS] before a switched word or phrase, while [LT] was used for Lithuanian. To facilitate cross-referencing, speakers' turns are numbered within the extracts that are quoted and analysed in the thesis.

Finally, it should be mentioned that a sample of the corpus, which consists of one transcribed recording is included in Appendix C. Ideally, one would want to append the whole corpus to enable the reader to refer directly to the recordings from which a given extract originates for further clarification or context. However, while in absolute terms the present corpus would be classified as quite small compared to other existing corpora, it is nevertheless too substantial in length to be included as a whole.

Chapter 4

Discourse Markers

The aim of this chapter is to assess the evidence for the first premise of the Emerging Mixed Code/Fusion Hypothesis, which states that if Russian discourse markers can appear as single switches in the Polish base without triggering further switching, have no competing equivalents from Polish and are exclusively employed in both Polish and Russian contexts, then this can be taken as one piece of evidence in favour of the emerging Polish-Russian mixed code. In order to test this premise, the analysis will focus on discourse markers.

As the existing studies demonstrate, discourse markers are often placed at the very top of the borrowability hierarchy in language contact situations, which can be attributed to a number of factors. They tend to perform pragmatic, rather than semantic functions and are often likened to function words. Much like function words, discourse markers are used with relatively great frequency and regularity, which means that they are well entrenched in the speakers' mental lexicons. Their salience is further reinforced by the fact that they are placed at the periphery of the clause, which makes them even more 'noticeable' in situations of language contact. Their clause-peripheral positioning could be linked to the fact that they function predominately on the discourse-pragmatic level of the linguistic system, where they are responsible for

what some researchers describe as ‘monitoring and directing’ of the hearer’s responses and reactions to the speaker’s utterances (Matras 2000: 505). The fact that they are unbound and do not enter into syntactic relations within the clause further contributes to their popularity in language contact situations.

As the existing literature suggests, by analysing the discourse and grammatical properties of discourse markers, as well as comparing their distribution against that of the native-language equivalents it is possible, to some extent, to situate a given instance of language alternation as a case of prototypical CS, CM or a fused lect. As will be demonstrated in this chapter, the first premise of the Emerging Mixed Code/Fusion Hypothesis is borne out by the Polish-Russian data. It will be argued that Russian discourse markers can not only appear in the Russian and Polish base without triggering further switching, but in almost all instances they have become the default option and have nearly replaced their Polish equivalents.

The chapter starts with a more detailed introduction into this broad category, as well as a survey of the existing research that deals specifically with the link between language alternation and discourse markers. It is then followed by a general overview of the Polish and Russian discourse markers which appear in the present data, with a focus on two Russian discourse particles, ‘nu’ — *well* and ‘vot’ — *here, there*, discourse markers ‘koroče’ — lit. ‘shorter’ — *in short*, ‘tipa’ — *sort of* and ‘prosto’ — *simply*, as well as Russian swearwords ‘blin’ — *damn* and ‘bljad” — *whore*, which tend to act as discourse-structuring devices, and thus perform similar functions to those of discourse markers.

4.1 The Category of Discourse Markers

When engaging in the conversational action, speakers may choose to provide information in addition to the propositional meaning which is already encoded in the utterance.

This can be achieved, among other things, with the help of set metatextual words and expressions, such as the ones shown in the examples below, taken from the present data:

Example 1

i ona ^[RUS] **nu** (.) ma klucze ona
and she well (.) has keys she

‘and she well (.) she has the keys’

Example 2

^[RUS] **vot** jęgo to ^[RUS] tochno by nie poznała
well him that definitely would not recognise

‘well him I would definitely not recognise’

The first remark that needs to be made with regard to these examples is that it is difficult to find an exact equivalent in English that would adequately convey their meaning. The highlighted words carry hardly any semantic content, and as such, they do not contribute to the propositional meaning of the sentence and could therefore be removed without causing any consequences for the syntactic or semantic make-up of the sentence. As the above examples demonstrate, these lexical items function within the realm of pragmatics. Consequently, an accurate English translation would be the one that conveys the pragmatic effect that these lexical items have in the context of the utterance in which they appear, hence *well* has been chosen as the closest contextual equivalent for both ‘nu’ and ‘vot’ in the two examples above.

Such metatextual words and expressions have been variously termed and defined in the existing literature. Following Schiffrin, they are often referred to as ‘discourse markers’ and are defined as “sequentially dependent elements that bracket units of talk” (1987: 31). One of the main problems and difficulties relating to the category of discourse markers is the great inventory of forms that it is said to include and the multitude of functions which they are supposed to fulfil. Some of the lexical items usually ascribed to this category include interjections (*oh*), adverbs (*well*), particles

(Russian ‘nu’ – *well*), verbs (*you know*), coordinate and subordinate conjunctions (*and, but, because*), and even entire sentence fragments and clauses (*do you see what I mean*). What unites them under one class of lexical items is their “functional similarities and partially overlapping distributions” (Schiffrin 1987: 65), i.e. these items share some of the basic features, but rarely (if ever) all of them.

As Jucker and Ziv put it (1998: 2), the membership in the class of discourse markers can be regarded as ‘scalar’, i.e. some lexical items demonstrate more of the ‘criterial features’ and can therefore be regarded as more prototypical members of the class, while other items display fewer characteristic properties and can be regarded as more peripheral. Perhaps two of the most ‘criterial features’ that are shared by all members of this class are based on their semantic and syntactic properties. From the semantic point of view, discourse markers are regarded as not contributing to the propositional meaning, while from the syntactic point of view, discourse markers are described as occurring outside the syntax of the utterance, or very loosely attached to it, and thus having no well-defined grammatical function (Brinton 1996: 32–35). It should be noted that Brinton’s definition is based on a strong division between semantics and pragmatics, a position which has been strongly criticised by the adherents of the so-called ‘reductionist approach’. However, despite its apparent limitations, it offers a good starting point for navigating through a number of different classifications and descriptions that have been proposed in the existing literature since it focuses on the core necessary features of lexical items that can be labelled as discourse markers.

One way of dealing with the wide variety of forms subsumed under one category is to break them down into smaller groups of structurally related items, which to some extent has also been reflected in the variety of terminology used in the literature. According to Jucker and Ziv (1998: 2) the term ‘discourse marker’ is convenient because it has “the least restricted range of application”, thus offering a possibility to include a broad inventory of lexical forms under one label. On the other hand,

Fischer (2006: 4) makes a distinction between a ‘discourse marker’ and a ‘discourse particle’. She describes the former as “the most inclusive”, while the latter as a “more restrictive term” (2006: 4).

Fischer (2006: 15) further claims that the definition of discourse markers and particles may be approached from a formal and functional side. From a functional point of view, these items “are taken to fulfil discourse functions, which are understood to be a subset of pragmatic functions in general” (ibid). On the formal side, the distinction can be made between ‘lexicalised’ and ‘non-lexicalised’ items. She concludes that on the basis of this typology, “discourse particles are items that are both formally and functionally defined, while discourse markers may be both lexicalised and non-lexicalised items that fulfil discourse functions” (2006: 15). Following this line of reasoning discourse particles are therefore lexicalised items that perform discourse-related functions, and which have the additional grammatical property of being particles. Discourse markers, on the other hand, is a much broader term used for any item fulfilling discourse functions, lexicalised or otherwise (thus including particles).

While Fischer’s definition offers a good starting point in distinguishing between discourse markers and particles, it can be expanded even further. On the basis of the discussion presented above, the following working definition can be formulated:

Discourse markers can be defined as lexicalised or non-lexicalised items that lack a precise semantic definition and perform a discourse-related function, e.g. marking discourse boundaries, relinquishing/taking the floor, or initiating clarification or repair. From a syntactic point of view, they occur outside the syntax of the utterance, and thus have no well-defined grammatical function. They are often found in clause-initial positions, though they may also appear in medial and final positions.

Within the broad category of discourse markers, discourse particles can be distinguished as a separate subcategory. They share all the prototypical

features of discourse markers, but have the additional grammatical property of being particles and can therefore be grouped together into a separate subcategory.

As far as the present data is concerned, the discussion in this chapter will focus on such Russian discourse markers as ‘koroče’ — lit. ‘shorter’ — *in short*, ‘tipa’ — *sort of* and ‘prosto’ — *simply*, as well as the discourse particles ‘nu’ — *well* and ‘vot’ — *here, there*. In the most general sense, all of these lexical items will be referred to as ‘discourse markers’. However, the term ‘discourse particle’, which as Fischer demonstrates is included under the broader label ‘discourse marker’, would be a much more precise description for ‘nu’ — *well* and ‘vot’ — *here, there*, and will therefore be employed in the present chapter.

4.2 Discourse Markers and Code-Switching/Mixing

The correlation between CS and discourse markers has already been addressed in the existing literature. One of the first remarks on this subject can be found in Poplack’s study of Spanish-English CS in the Puerto-Rican community in New York (1980). Although her work is better known for focusing on the syntactic constraints of switching, Poplack does mention that discourse markers, tags, interjections, and fillers make-up 29% of all instances of language alternation in her corpus (1980: 602). Poplack’s (1980: 605–608) interpretation of the results that she obtained mainly focuses on the fact that such elements do not require much proficiency in English, which means that speakers whose dominant language is Spanish can insert them with ease to demonstrate their bilingual competence. She further adds that such switches can be regarded as ‘emblematic’ (1980: 614), as they ensure that English, which can be regarded as a marker of prestige among working-class Puerto Ricans is present in predominately Spanish conversations.

But, as Matras rightly points out, “different communities may have different motivation for switching, which in turn might be reflected in different patterns of switching behaviour” (1998: 288). While in the Puerto Rican community English may well be a symbol of prestige and is inserted into Spanish conversations with an aim of projecting higher social status, it seems to serve a different purpose in other situations of language contact. In her study of Israeli–English discourse, Maschler (1994, 1998) goes beyond the social implications of language alternation at discourse markers, but rather focuses on the communicative effect that it achieves. She found that Hebrew discourse markers are inserted into Hebrew-English discourse in order to ‘frame’ and mark conversational action boundaries (1998: 136). As Maschler points out, switching at discourse markers further reinforces their discourse framing and structuring function. Thus, Maschler’s explanation focuses mainly on the relationship between language alternation and the actual communicative function that is normally attributed to these lexical items, rather than the social implications of their use.

Matras (1998, 2000) offers yet another view on language switching at discourse markers, centred around the notion of ‘fusion’. In communities where bilingualism has become a norm, speakers may over time adopt just one set of particular linguistic structures, thus entirely eliminating or partially reducing the need to select among competing word-forms. This is usually the case when discourse marking items are involved, which in Matras’ account include conjunctions, tags, interjections, and focus particles (1998: 144–145). According to Matras (2000: 516), discourse markers are elements that help speakers monitor and direct the way a propositional content of the utterance is processed and accepted by the hearer. The monitoring and directing operations are cognitively more complex than formulating propositional content, for the latter involve back-processing, planning ahead, anticipating and controlling reactions, interpreting gestures and intervening with a possible undesired course of processing on the part of the speaker (Matras 2000: 517). Therefore, bilingual speakers will aim to reduce this cognitive load by eliminating the choice between

the two systems available at their disposal. In doing so, one set of discourse-marking devices is usually abandoned at the expense of the other.

Matras (*ibid*) further points out that ‘fusion’ may have both synchronic and diachronic manifestations. In the case of synchronic manifestation, system A and system B variants of the same variable may be used side by side, with speakers alternating between the available forms. He exemplifies this with the case of Low German speakers in the United States who constantly alternate between Low German connectors ‘und’, ‘aber’ and their English counterparts *and*, *but* (2000: 524). The fact that English connectors frequently appear in Low German conversations suggests that speakers have accepted them as part of the available inventory of forms, thus partially eliminating the need to maintain the separation between the two systems. Although full replacement of one system by another has not taken place yet, it is nevertheless ‘synchronically licensed’, i.e. speakers have abolished the requirement to maintain a consistent separation between the Low German and the English system (Matras 2000: 525). It is of course possible that over time English forms will take over at the expense of their Low German counterparts, leading to a complete replacement, which would then be treated as a ‘diachronic manifestation of fusion’.

As mentioned in the literature review in Chapter 2, Auer explicitly states that there is a direct link between switching at discourse markers and the emergence of a mixed code. In characterising a middle position on the continuum between CS and CM, Auer (1998: 16–21) states that one of the steps involves the frequent use of (nonce) borrowings, which are often content words, in particular nouns. He further adds that “other single-word borrowings that often contribute to the formation of mixed codes are language B discourse markers, which, in many bilingual communities, occur in conversations in language A” (1998: 17). This idea is further elaborated in his later work on emerging fused lects, which is an even further development from language mixing, when the use of one language or the other for certain lexical items becomes

obligatory (1999: 321–323). The transition from language mixing to a fused lect starts at what Auer calls ‘relatively unbound elements of grammar’ (1999: 324), which include discourse markers. These unbound elements are usually affected by structural sedimentation sooner and faster than other elements, which suggests that the transition from language mixing to fused lects for the unbound elements does not usually coincide with the transition for the other structural domains (*ibid*). This regularity is largely conditioned by psycholinguistic factors. In the case of discourse markers an important role is played by the fact that they are highly recurrent and occur rather frequently in spoken discourse. Each time a discourse marker is evoked in language production, its level of entrenchment in the mental lexicon is reinforced, hence there is a direct link between the frequency with which a speaker encounters a given lexical unit and its level of cognitive entrenchment (Ziem 2014: 296). Since discourse markers constitute a set of forms that are employed with relatively great frequency and regularity, their level of entrenchment is likely to be high, which makes them not only particularly productive, but also explains why they are often placed at the top of the borrowing hierarchies.

Finally, the correlation between language mixing and discourse markers has also received some attention in Clyne’s account (1991, 2003) of switching prompted by so-called ‘trigger words’. Trigger words can be defined as lexical items that are “at the intersection of two language systems, which, consequently, may cause speakers to lose their linguistic bearings and continue the sentence in the other language” (Clyne 1991: 193). As Clyne further explains in his later work, trigger words simultaneously belong to more than one language (from the speaker’s and/or from the speech community’s point of view), and may facilitate the transition from one language to another (2003: 162). It should be noted that trigger words themselves should not be regarded as switched items because they are already part of the language of interaction (Riehl 2005: 1946). Switching is facilitated by the mere fact that a trigger word is at the same time part of the ‘other’ language and can therefore attract more ‘other’ language

material, though it may also be employed as a single switch without any triggering effects (ibid).

In Clyne's classification lexical transfers constitute one of the largest categories of trigger words (2003: 162). They are defined as lexical items that are phonologically unintegrated or integrated at a low level only in the language of interaction. Another closely-related category of trigger words is that of discourse markers, which in Clyne's (1991: 194) account are categorised as a sub-group of bilingual homophones (words that sound the same or nearly the same in the two (closely-related) languages of the speaker). However, Riehl (2005: 1949) reports that discourse markers are so productive as trigger words that they merit to be included under a separate category. The fact that discourse markers can be classified as trigger words entails that on the one hand, they are (to a lesser or greater degree) integrated into the recipient language system, but at the same time they still belong to the system of the donor language. This would explain why they appear as integral elements of one system, but can also be found in multiple switches or even longer stretches of discourse in the other language.

4.3 Polish and Russian Discourse Markers in the Present Corpus

Since discourse markers constitute a very broad category, made up of a whole range of seemingly distinct lexical items, it seems reasonable to present a general overview of the inventory of both Polish and Russian forms that appear in the present corpus. Overall, there are 5,268 Russian discourse markers, as well as lexical items that function as discourse markers, inserted into the Polish base, which in turn means that these lexical items make up 58% of all the switches in the present data.

The first observation that needs to be made is that the inventory of Russian lexical items that function as discourse markers is much broader compared to the inventory of Polish forms. There are ten Russian discourse markers that appear with various frequencies throughout the corpus. They include such sentence adverbials as ‘koroče’ – lit. ‘shorter’ – *in short*, ‘tipa’ – *sort of*, ‘prosto’ – *simply*, ‘voobše/v obšem’ – *in general*, ‘kстати’ – *by the way*, ‘vrode’ – *kind of*, the particles ‘nu’ – *well*, ‘vot’ – *here, there*, and ‘tak’ – *so*, as well as the agreement particle ‘da’ – *yes*, which in many instances performs discourse marking functions. In contrast, there are only four Polish forms that can be identified as discourse markers, namely ‘niby’ – *kind of*, ‘jakby’ – *like*, as well as ‘to’ – *so* and ‘no’ – *well, okay*.

The quantitative discrepancy between the Polish and the Russian forms could partially be conditioned by the inherent differences between the two languages. Judging from the existing research, the category of discourse markers in Standard Polish, as defined in section 4.1, is not as productive as the equivalent category in Standard Russian. In the Polish linguistic tradition discourse markers are normally referred to as ‘operatory metatekstowe’ (*metatextual operators*). Wajszczuk (2005: 109–117) proposes a further subdivision into ‘metatextual’ and ‘metapredicative’ operators, where ‘metapredicative operators’ are defined as lexical items that function on the level of the sentence and that enter into syntactic relations with other constituents, while ‘metatextual operators’ are described as lexical items that often co-occur with other words, but do not form any syntactic relations. Examples of the former include such forms as adverbs that convey the meaning of intensity (e.g. ‘bardzo’ – *very*, ‘strasznie’ – *awfully*), approximation (‘trochę’ – *a little*, ‘prawie’ – *almost*) or adnumeration (‘niepełna’ – *hardly any*, ‘przeszło’ – *more than*); examples of the latter include such particles and conjunctions as ‘nawet’ – *even*, ‘też’ – *also*, ‘zwłaszcza’ – *especially*, ‘chyba’ – *maybe*, ‘na pewno’ – *definitely*, ‘i’ – *and* (ibid).

In contrast, the equivalent category in Russian linguistic tradition, the so-called

‘vvodnye slova i predloženija’ (*introductory words and phrases*), is defined rather broadly and consists of qualitatively different forms from the Polish ‘metatextual and metapredicative operators’. According to the Academy Grammar of Russian, ‘vvodnye slova’ are used to supply metatextual information that is additional to what is contained in the main body of a sentence, and thus in writing, they would normally be singled out by commas (Švedova 2005: 228). The metatextual information may include the speaker’s attitude towards what is being said, or additional characteristics of the message being expressed. From the point of view of grammatical organisation, this group is very diverse. It is said to include conjugated forms of the verb (e.g. ‘znaeš’ — *you know*), adverbs (e.g. ‘prosto’ — *simply*), adverbial participles, particles (e.g. ‘nu’ — *well* and ‘vot’ — *here, there*), as well as idioms (e.g. ‘vo vsjakom slučae’ — *in any case*).

The problem with Wajszczuk’s (2005) taxonomy is that lexical items classified under the categories of ‘metapredicative operators’ and ‘metatextual operators’ do not behave the same way as discourse markers, as defined in section 4.1. Such lexical items as ‘bardzo’ — *very* and ‘strasznie’ — *awfully* carry a clear semantic meaning and their deletion would alter the propositional content of the utterance. Similarly, most of the particles and all of the conjunctions classified under the subcategory of ‘metatextual operators’ cannot be detached from the utterance without changing its meaning or rendering it ungrammatical. In order to maintain uniformity and make a fair comparison between the Polish and Russian forms, the same syntactic and semantic criteria will be applied. Therefore, a lexical item will be classified as a discourse marker if it carries no semantic meaning and can be removed from the sentence without affecting its propositional content or syntactic make-up. Following these criteria, the inventory of Polish forms is much more limited compared to the available Russian forms.

The second observation that needs to be made with regard to Polish discourse markers

is that they are highly idiolectal. For example, the discourse marker ‘niby’ — *kind of* is employed on a total of 30 occasions by speakers C, R and K. However, in the discourse of the former two speakers it is employed in one instance only, which means that the remaining 28 instances belong to speaker K. Likewise, there are seven instances of the Polish discourse particle ‘no’ — *well*, two instances in the discourse of speaker B, and five instances in the discourse of speaker J. This shows that Polish discourse markers are not only restricted to the discourse of selected speakers, but are also employed with very limited frequency.

By comparison, the Russian ‘koroče’ — *in short* is employed on 306 occasions by nineteen speakers, ‘tipa’ — *kind of* is equally popular with a total of 364 instances of use across all the speakers in the data, while ‘prosto’ — *simply* is used on 41 occasions by eleven different speakers. However, the two most popular discourse markers are undoubtedly the discourse particles ‘nu’ — *well*, ‘vot’ — *here, there*, with 2,724 tokens of the former, and 710 tokens of the latter. This demonstrates not only that there is a significant quantitative difference between the different types of Polish and Russian discourse markers in the present corpus, but also that the Russian discourse markers clearly outnumber the Polish ones.

In addition to discourse markers and discourse particles, it is also worth focusing on swearwords which function as discourse-structuring devices. Two of these lexical items regularly feature in the present corpus, namely the more vulgar ‘bljad’ — *whore*, also employed in its abbreviated version ‘blja’, as well as its more polite equivalent ‘blin’ — *damn*. ‘Blja/bljad’ — *whore* is used by 22 speakers on 719 occasions, while ‘blin’ — *damn* appears in the discourse of sixteen speakers with a total number of 301 occurrences. While ‘bljad’ can be employed to convey its literal meaning *whore*, there are no such instances in the corpus. Instead, its use is comparable to that of *fuck* and *shit* in English, i.e. it functions as an emotionally charged discourse-structuring device.

The Polish equivalent of ‘blja/bljad’ is the swearword ‘kurwa’, which can also be employed in its literal meaning of *whore*, or as an emotionally charged expletive, which can also perform discourse-related functions. What is significant is that there is only a single instance of ‘kurwa’ — *whore* in the present data, where it is used in its literal meaning, i.e. the speaker refers to someone as a whore. This finding suggests that the Russian swearwords ‘blin’ — *damn* and ‘bljad’ — *whore* are the default option that the speakers exclusively resort to. Since these lexical items contribute to the organisation of the interaction and the structuring of verbal exchange they will be analysed together with discourse markers and discourse particles in the present chapter.

4.4 The Discourse Particles ‘nu’ and ‘vot’

The two discourse particles that regularly appear throughout the corpus are Russian ‘nu’ — *well* and ‘vot’ — *here, there*, which according to Zemskaya (1979: 90), are also two of the most frequently used discourse particles in the spoken Russian language. This comment can also be applied to the present corpus, where there are as many as 2,724 instances of ‘nu’ and 710 instances of ‘vot’. Unlike other discourse markers, these two particles are employed in both Polish and Russian contexts by all the speakers in the data, further attesting to their even distribution across the corpus and acceptance within the Polish speech community.

The observations presented above suggest that ‘nu’ and ‘vot’ are very strong candidates for the hypothesis of ‘fusion’ in what Matras calls its ‘diachronic manifestation’, which can be formulated as follows: if ‘nu’ and ‘vot’ can appear as single switches in a Polish base, and have no competing equivalents from Polish, then they can be regarded as one piece of evidence in favour of the Emerging Mixed Code hypothesis. In order to test this hypothesis further, the following questions will be addressed:

(1) What are the Polish equivalents of ‘nu’ and ‘vot’, and are they attested in the present data? (2) Have ‘nu’ and ‘vot’ been ‘fused’ along with the majority of their corresponding functions (as presented in table 4.1)? (3) Are they able to appear as stand-alone items (i.e. single switches) in a Polish base?

A general overview of discourse-related functions performed by ‘nu’, ‘vot’ and the combination ‘nu vot’, as well as the exact number of occurrences for each function is shown in Table 4.1. It should be borne in mind that such a classification inevitably carries a certain degree of vagueness, for it is based on the subjective interpretation of the researcher whose judgement (at least in certain cases) may be different from someone else’s. The purpose of the table is to compile a taxonomy of functions that have been replicated in the present data and to compare them with the ones that ‘nu’ and ‘vot’ are said to perform in Standard Russian in order to establish the degree of integration of these particles in the Polish-Russian mixed code. In the majority of cases, the functions attested in the present corpus correspond to the ones that are listed in other sources, namely Gorbachevich’s (2004) ‘Большой академический словарь русского языка’ (*Large Academic Dictionary of the Russian Language*) and Zemskaya’s monograph on spoken Russian (1979: 88–93).

Functions	[RUS] nu	[RUS] vot	[RUS] nu vot	[PL] no
Beginning a question or request	60	1		
Beginning an answer or response	403		2	
Beginning a new topic or subtopic	335	8	18	
Ending a turn	45	1		
Beginning direct speech	65	6	2	
Beginning new turn	114	5	1	
Searching for the required word or hesitating	183	37	1	4
Emphatic function (often with other discourse markers/conjunctions)	402	250	5	3
Demonstrative function		312	6	
Expressing agreement	675			
Emphatic agreement (nu + да)	146			
Expressing indication that the listener is following what is being said	183	2		
Expressing generalisation or conclusion	29	86	14	
Resuming the conversation after a digression	28	18	14	
Introducing a counter argument	56		1	
TOTAL	2724	710	62	7

Table 4.1: Russian ‘nu’, ‘vot’, ‘nu vot’ and Polish ‘no’: functions and numbers

4.4.1 Discourse Particle ‘nu’

With a total number of 2,724 instances in the corpus, ‘nu’ is undoubtedly the most popular discourse particle in the present data. It is employed on numerous occasions by all the speakers, which is a good starting point for the hypothesis of ‘fusion’. The closest partial equivalent from Polish that covers some of the functions attributed to ‘nu’ is the discourse particle ‘no’. Like its Russian counterpart, the Polish ‘no’ is often employed in turn-initial positions, where it precedes a question or response, or indicates the beginning of a new topic, subtopic or direct speech. It may also help to create structure within conversational exchange by indicating to the listener that the speaker is searching for the right word or hesitating, or simply be employed to reassure the speaker that the listener is following what is being said. Moreover, ‘no’ is also said to perform an emphatic function, especially when it occurs with other discourse markers, conjunctions or parenthetical verbs. As stated in Doroszewski’s (2011) ‘Słownik Języka Polskiego’ — *Polish Language Dictionary*, similarly to the Russian ‘nu’, the Polish ‘no’ can also be used as a particle expressing agreement, in which case it would be translated into English as *yes*.

The corresponding Polish discourse particle ‘no’ can be found in the discourse of two speakers in the corpus, B and J, where it co-occurs with the Russian ‘nu’. The Polish ‘no’ is employed on seven occasions (two for speaker B and five for speaker J), while the Russian ‘nu’ is used on 166 instances by speaker B and 90 instances by speaker J. Such significant quantitative discrepancy leaves little doubt as to the fact that the Russian ‘nu’ is the default option for both speakers, as well as all other speakers that feature in the present corpus. To put it in quantitative terms, the Polish ‘no’ makes up only 0,26% of all instances of ‘no’/‘nu’ in the corpus, while the Russian ‘nu’ amounts to the remaining 99,74%. This further corroborates the view that it is the Russian ‘nu’ which has become the default option that the speakers resort to.

If ‘nu’ is to be regarded as the default option, it would be expected to appear in

stretches of Russian discourse, as well as the Polish base. While the Polish ‘no’ is limited only to the Polish base, the Russian ‘nu’ can in fact appear in both settings. In 1,961 occurrences (72%) ‘nu’ either constitutes a turn on its own, or features as a single switch in the Polish base. In the remaining 763 instances (28%) it appears with other Russian lexical items in the initial, medial or final position as part of a larger stretch (i.e. as a multiple-word switch), ranging from several words to the entire turn. The fact that in the majority of cases ‘nu’ is actually embedded on its own into the Polish base is significant. It suggests that the Russian discourse particle has become fully integrated into the recipient language system and is now regarded as the default option. This can also be regarded as an important piece of evidence in favour of the hypothesis of ‘fusion’.

As far as functions and meaning are concerned, the Russian ‘nu’ is more prominent than the Polish ‘no’. Despite a very limited number of occurrences it is evident that ‘no’ does not display its whole range of functions in the Polish-Russian data. In speaker J’s discourse there are five instances of ‘no’, four of which display the same function, namely that of searching for the right word, hesitating or trying to win some time before formulating a thought. An example of such usage can be found in Extract 1:

Extract 1

K is asking J why she thinks her project will be successful.

- (1) K: i w jaki to sposób musi być— czemu ty myślisz że ^[RUS] **vot**—
 and in what it manner must be— why you think that well—
 ‘and how is this supposed to be— why do you think that well’—

^[RUS] **babki** dadzo? bo to jakoś ^[RUS] **nu** gry ^[RUS] **nu** nie wiem
 money give? because it somehow well games well not know

‘[they will] give money? because it somehow, well, games, well, I don’t know’

- (2) J: no:: (.) bo my— bo napisalam tak ze—
 well:: (.) because we— because wrote so that—
 ‘well (.) because we— because I wrote so that’—

J begins her response in turn (2) turn with ‘no’ characterised by a slightly longer pronunciation of the vowel /o/, which can be regarded as an additional means of conveying hesitation. Another piece of evidence which shows that the speaker is hesitating and needs more time to formulate her reply is a false start that she makes right after ‘no::’. When she finally succeeds in formulating the first part of her response, she gets interrupted by K and therefore does not get to finish her turn.

What is significant is that the same speaker also employs the Russian ‘nu’ in exactly the same context with exactly the same function:

Extract 2

K and J continue talking about the project.

- (1) K: aj a co tam trzeba bardzo mocne sponsorowanie czy co tam?
 oh and what there need very strong sponsoring or what there?
 ‘oh and does one need very substantial sponsoring there or what?’
- (2) J: _[RUS] nu:: (.) mhh (.) jezeli chcesz kupic gry planszowe dobre—
 well:: (.) mhh (.) if want buy games board good—
 ‘well (.) mhh (.) if you want to buy good board games’—

Just like ‘no::’ in Extract 1, the turn-initial ‘nu::’ is pronounced with a slightly longer vowel /u/ and is followed by ‘mhh’, which further shows that the speaker is hesitating and needs some extra time to formulate her response. As far as this particular function is concerned, speaker J is using ‘nu’ and ‘no’ interchangeably. On the basis of the recording which features in the corpus it can be concluded that the speaker allows for some degree of overlap with regard to only this one function, as the Russian ‘nu’ is chosen as the default option in all other instances.

In speaker B's discourse 'no' appears in only two instances where its function can be described as 'emphatic', for in both instances it is followed by another discourse marker 'to'. In the first instance it appears in the beginning of a response ('no to dobrze' – *well that's fine*), while in the second instance in the middle of the turn with another particle and a parenthetical verb ('no to wiesz' – *so you know*). This means that in the present corpus 'no' performs only two functions, either emphatic or to indicate hesitation on the speaker's part. However, the main purpose of the overview of 'no' is not so much about establishing the functions that it performs, but to further reiterate the fact that compared to 'nu', its presence in the corpus is very negligible, both from a functional and quantitative point of view. While 'no' is limited to only seven instances with an extremely narrow range of functions, its Russian equivalent dominates not only in the discourse of J and B, but also all other speakers in the corpus where it displays the full range of function typically attributed to 'nu' in Russian lexicography, e.g. Gorbachevich's (2004) dictionary. This suggests that a great majority of functions that 'nu' is said to perform in Standard Russian has been replicated in the emerging Polish-Russian mixed code, and further attests to its status of a fully integrated lexical item.

To some extent, the functions that 'nu' performs in the data depend on the position in which it appears within a sentence. When used in an utterance-initial position it often introduces a question or request, or the speaker's comment or response to the interlocutor's previous turn. An example of this type of usage can be found in Extract 3:

Extract 3

- (1) K: wiesz gdzie tam obok opuszczony _[RUS] **komissariat** jest?
 know where there next to abandoned police station is?
 'do you know where the abandoned police station is there?'

- (2) L: _[RUS] **nu** slyszal ale ja nie byl tam _[RUS] **ni razu**
 well heard but I not been there not once
 ‘well I’ve heard [about it] but I haven’t been there’

The initial ‘nu’ appears in the beginning of turn (2) where it introduces the answer to speaker K’s question. Although there is a cluster of two Russian lexical items at the end of L’s turn, the material that immediately follows the Russian discourse particle is in fact Polish. This suggests that there is no evidence to assume that the insertion of ‘nu’ determines the base language for the rest of the utterance, since it is expressed in Polish. Neither is there any evidence to suggest that there might have been a trigger in the preceding turn that prompted the speaker to employ a Russian lexical item. Apart from one single Russian switch ‘komissariat’, turn (1) is also expressed in Polish, which supports the view that ‘nu’ can be regarded as the default choice that is part of the speaker’s lexicon.

When it appears in the middle of a sentence, ‘nu’ can indicate that the speaker is about to provide a more detailed explanation or precision, that a speaker is searching for the right word or hesitating, or simply performs an emphatic function. In the latter case ‘nu’ often appears with another discourse marker/particle or conjunction. Some examples can be found in Extracts 4 and 5:

Extract 4

- (1) N: to oni wtedy do końca miesiąca jak przepracował wtedy oni mnie
 so they then till end month when worked then they me
 ‘so they then, [when] I worked until the end of the month, then they’
- zapłacili (.) _[RUS] **nu** (.) piętnastego tego
 paid (.) well (.) fifteenth this
 ‘paid me (.) well (.) [on the] fifteenth’

N employs ‘nu’ in the middle of his utterance before he provides a more detailed explanation of the exact day when he was paid. It occurs in an entirely Polish base

without any identifiable triggers that could be held responsible for its insertion, and without having any consequence on the language of the utterance.

Another example can be found in Extract 5:

Extract 5

- (1) G: u mnie ^[RUS] **nu** (.) jakby wszystko i po staremu
at me well (.) seemingly everything and according to old
'at my end well (.) everything seems to be as before'

In the above example 'nu' is followed by a brief pause in order to indicate that the speaker may be searching for the right word, or is simply taking his time to formulate the turn. As with the previous examples, 'nu' appears on its own in an utterance with a Polish base.

There are not many examples of 'nu' that occur in the end of an entire turn, but it can often be found in utterance-final positions within a turn. An example can be found in Extract 6:

Extract 6

- (1) C: to ^[RUS] **beton** taki?
it concrete such?
'is this [a kind of] concrete?'

- (2) D: ^[RUS] **da**
yes
'yes'

- (3) C: zalewajo ^[RUS] **tipa**?
pour kind of?
'[do they] kind of pour [it]?'

- (4) D: ja tobie mówia ^[RUS] **nu** (.) taki aż czarny ^[RUS] **beton blja**
I you tell well (.) such even black concrete shit

‘I am telling you (.) such black concrete, shit’

In the above extract the emphatic ‘nu’ is used at the end of the first utterance in turn (4) in order to reinforce the proposition that D wishes to express. In addition to performing an emphatic function, the discourse particle also functions as an utterance-closing device. As with previous examples, it is clearly acceptable to employ the Russian discourse particle ‘nu’ in a Polish base.

One of the functions of ‘nu’ that is most frequently attested in the present data is when it appears as the only lexical element within a turn in order to express agreement or to indicate that the listener is following what is being said. An example of this usage can be found in Extract 7:

Extract 7

N and M are talking about N’s payday.

- (1) N: ja teraz poczekaj (.) piętnastego dostawszy był za p-p-p- za pięć dni
I now wait (.) fifteenth received was for p-p-p- for five days
‘I now — wait (.) [on the] fifteenth received for f-f-f- for five days’
- (2) M: _[RUS] **nu**
O.K.
O.K.
- (3) N: a teraz _[RUS] **dolžen vo na dnjah** znowu dostać
and now must here on days again receive
‘and now I must receive [the payment] again these days’

The occurrence of ‘nu’ in turn (2) acts as an indication to the speaker that the listener is following what is being said and an encouragement for N to carry on. As Table 4.1 demonstrates, it is one of the more frequent functions of ‘nu’ in the present corpus, with a total number of 183 occurrences.

Another very frequent function that ‘nu’ performs when it is used on its own is that of an agreement or confirmation marker. An example can be found in Extract 8:

Extract 8

- (1) G: a tak w szkole po ile lekcji?
and like this in school by how many lessons
‘and normally at school how many lessons [do you have]?’
- (2) F: po osiem
by eight
‘eight’
- (3) G: każdy dzień po osiem?
every day by eight
‘every day eight?’
- (4) F: _[RUS] **nu**
yes
‘yes’

The main function of ‘nu’ in turn (4) is to provide an affirmative response to G’s question in turn (3), with the meaning of ‘yes’. This can be described as a shorter version of a more emphatic affirmation that also frequently features in the present data, namely ‘nu da’:

Extract 9

- (1) G: _[RUS] **avtomekhanika** musi lepiej?
car mechanics must better
‘car mechanics must [be] better?’
- (2) F: _[RUS] **nu da**
well yes

The main aim of the above discussion was to test the Russian discourse particle ‘nu’ for the hypothesis of ‘fusion’ in what Matras refers to as its ‘diachronic form’. The mere fact that ‘nu’ can be inserted on its own in a Polish base and that it constitutes the default choice for virtually all the speakers in the corpus suggests that it is reasonable to treat it as a discourse particle that has been ‘fused’ into the Polish-Russian mixed code. The positions within a turn where ‘nu’ can appear, as well as the range of functions that it can perform are in fact compatible with the ones that are characteristic of ‘nu’ when it is employed in Standard Russian. It suggests that the particle has been ‘fused’ into the emerging Polish-Russian mixed code with the majority of its corresponding functions.

While the Polish ‘no’ does appear in the corpus, from the quantitative view its presence can be described as negligible, for it features in the discourse of two speakers only with a very limited range of functions. However, the very fact that it is present might suggest that it is still too early to claim that the data display a diachronic manifestation of fusion and that the possibility of it being a synchronic manifestation should be kept open. Moreover, its presence can be regarded as a further indicator that the Russian ‘nu’ should really be treated as such, rather than the local Polish reduction of ‘no’ to ‘nu’. The Russian particle is clearly the default option for all the speakers and it can be employed to perform the full spectrum of functions. This finding suggests that the replacement of ‘no’ with ‘nu’ is virtually complete, and the data display a diachronic manifestation of the process that has already taken place.

4.4.2 Discourse Particle ‘vot’

With a total number of 710 occurrences, ‘vot’ is the second most popular discourse particle that appears in the corpus. It seems to display some of the same characteristics as ‘nu’: it can occur in stretches of both Russian and Polish discourse, and can appear

on its own in a Polish base. Unlike ‘nu’, the discourse particle ‘vot’ does not have any cognates or directly corresponding equivalents in Standard Polish. On the basis of these observations it could be argued that ‘vot’ has been borrowed into Polish and should be regarded as a Polish discourse marker. However, this would not be accurate for several reasons. As will be demonstrated in the remainder of this section, this discourse particle has been borrowed into the Polish-Russian mixed code. In order to label ‘vot’ as ‘Polish’, it would have to be borrowed into the regional variety of Polish as well, which is still spoken by some speakers in Vilnius and the Vilnius region. However, this cannot be verified with the data collected for the present corpus. It may well be the case that ‘vot’ would still be employed when speakers communicate in the regional variety of Polish without switching into Russian, but this could only be checked against the data that document the regional variety of Polish, and not the Polish-Russian mixed code.

While ‘vot’ does not have a directly corresponding equivalent in Standard Polish, there are other lexical items that cover some of its functions. The closest equivalent in Standard Polish would be the particle ‘oto’, but it does not appear in the present corpus. Another particle that could be regarded as a partial equivalent of ‘vot’ is the Polish ‘ot’, which is employed in the data on a total of eighteen occasions by six different speakers. However, there is a considerable amount of ambiguity surrounding this particle, which can be attributed to several reasons. The first reason is that phonologically, ‘vot’ and ‘ot’ are extremely close, which means that the instances that were counted as ‘ot’ might actually be instances of ‘vot’ where the initial /v/ has not been articulated clearly enough. The second ambiguity comes from the fact that there is a homophonous Russian particle ‘ot’, and in the variety of Polish spoken in Lithuania the pronunciation of the Polish ‘ot’ and the Russian ‘ot’ is virtually the same. Therefore, as far as ‘ot’ is concerned, it appears that the distinction between the Polish and the Russian particle can only be made on the basis of semantic criteria.

According to Dubisz's (2006) dictionary, 'ot' is defined as a particle that may be employed before various constituents of a sentence in order to show that the speaker regards something as insignificant, ordinary or trivial, e.g.: 'Dostał za to, ot, parę złotych'. While it is problematic to find an exact English equivalent, the example can be translated as *He received for it only a few zlotys*. According to Gorbačevič's (2004) dictionary, the Russian 'ot', on the other hand, can be employed to express two meanings: 1) to emphasise and single out the word that it is attached to, e.g.: 'Hleb-ot, umeeš' est'?' — *Bread, do you know how to eat [it]?* 2) to reinforce or emphasise the word that follows after it (synonymous with 'vot'), e.g.: 'Prišel ot pogljadet'.' — *I came well, to have a look*.

Upon a closer analysis of the examples of 'ot' in the present data it becomes clear that they resemble the Russian patterns of use, rather than the Polish ones. There is not a single instance where 'ot' would be employed to convey that something is insignificant or trivial. On the contrary, it seems to perform a function very similar to that of 'vot' when the latter is used to emphasise or reinforce the meaning of the lexical item (or items) which immediately follows 'ot'. Some examples can be found in Extracts 10 and 11:

Extract 10

- (1) L: ja tobie pokaža (.) _[RUS] **ot** będzim wychodzić—
 I you show (.) *emph. particle* will be leave—
 'I'll show you (.) when we'll be leaving—'

In the above extract 'ot' appears after a brief pause and serves to emphasise the propositional content that follows right after it. In this sense, it is synonymous with 'vot' and could have been easily replaced by the latter.

Extract 11

to indicate that the speaker is searching for the right word, or ‘vot’ used to introduce a counter-argument or opposition. This considerable overlap in functions which ‘vot’ is said to perform in Standard Russian, and the ones which feature in the present data suggests that just like ‘nu’, it has been replicated in the emerging Polish-Russian mixed code with the majority of its corresponding Russian functions. Examples of the demonstrative, emphatic and concluding ‘vot’ can be found in Extracts 12–14:

Extract 12

- (1) L: (...) a patrza:j_[RUS] **vot** tam_[RUS] **vot** drzwi
 (...) and look there there there door
 ‘(...) and look over there there [is] a door’

Speaker L’s turn contains two instances of ‘vot’, the first one could be classified as demonstrative, for it appears right after ‘patrza:j’ in the imperative form and is followed by ‘tam’, which clearly indicates that the speaker is pointing the listener’s attention towards something. The second instance of ‘vot’ is slightly more ambiguous because it could modify either of the two lexical items, forming a cluster ‘vot tam vot’, or ‘vot drzwi’. In either case, it would be classified as emphatic. Despite the presence of ‘vot’, the base language of the utterance remains Polish, and the presence of the Russian discourse particle does not trigger any switching.

An example of the concluding ‘vot’ can be found in Extract 13:

Extract 13

A is talking about funeral arrangements.

- (1) A: (...) kupili sukienka (.) kupili pantofle (.)_[RUS] **nu vse** kupili
 (...) bought dress (.) bought shoes (.) well everything bought
 ‘(...) [they] bought the dress (.) bought the shoes (.) well [they] bought everything’
 wziali i wypisali ze szpitala (.) nie umarla ((laughter)) (0.4)
 took and discharged from hospital (.) not died ((laughter)) (0.4)

‘[they] took and discharged [her] from the hospital (.) [she] didn’t die ((laugh-
ter)) (0.4)’

(...) dobrze że grobu jeszcze nie kupili
(...) good that coffin still not bought
‘(...) good that [they] didn’t buy a coffin’

(2) B: _[RUS] **nu da**
well yes
‘well yes’

(3) A: to byłoby _[RUS] **voobše** (0.2) _[RUS] **vot**
it would be at all (0.2) there
‘now that would be quite something (0.2) there’

Speaker A concludes her story with ‘vot’ which occurs in the final position at the very end of turn (3). According to Zemskaya (1979: 91), this type of ‘vot’ acts as a concluding remark to what has been previously expressed by the speaker and normally forms a separate intonation unit. This observation can also be extended to the above example where the Russian discourse marker appears by itself after a brief pause, and constitutes a separate prosodic unit with a falling intonation contour. Although in the above extract ‘vot’ appears right after ‘voobše’, this does not mean that the appearance of the latter has been triggered by the presence of the former, for as the previous examples suggest ‘vot’ can be embedded into the Polish base on its own.

Finally, one of the less popular functions of ‘vot’ can be found in Extract 14:

Extract 14

N and M are talking about setting up direct debit.

(1) N: tam trzeba było numer karteczki gdzie na karteczce samej napisany
there must was number card where on card itself written
‘a card number the one [that is] written on the card itself was required’

- (2) M: _[RUS] **nu** ale bez róż— bez różnicy
 well but without diff— without difference
 ‘yes but it makes no diff— no difference’
- (3) N: a:: _[RUS] **vot** różnica była
 and there difference was
 ‘and yet there was a difference’

In turn (3) speaker N wishes to emphasise that there was a difference after all, which he partially achieves with the help of ‘vot’. In Zemskaya’s classification (1979: 91), this would be an example of ‘vot’ used to introduce a counter-argument or an opposition to what the interlocutor has been saying.

It seems that the Russian discourse particle ‘vot’ is an even better candidate for the hypothesis of ‘fusion’ than ‘nu’ since it does not have any immediate competing forms from Polish. Just like ‘nu’, ‘vot’ appears in either the initial, medial or final position of a sentence, and performs an equally broad range of functions, which can be likened to those that ‘vot’ performs in Standard Russian. Moreover, it appears as a stand-alone item in the Polish base, which means that the speakers have accepted it as part of their lexicon. These observations can be treated as evidence of the hypothesis of ‘fusion’ in its diachronic manifestation, which could also be shown to include the Russian discourse particle ‘nu’. Since discourse particles are essentially a subtype of discourse markers, it seems plausible to extend this hypothesis even further to include such forms as ‘koroče’ — lit. ‘shorter’ — *in short*, ‘tipa’ — *sort of* and ‘prosto’ — *simply*, which will be analysed in greater detail in the following subsection.

4.5 Discourse Markers ‘koroče’, ‘tipa’ and ‘prosto’

In addition to the two discourse particles analysed in the previous section, other Russian discourse markers that feature regularly in the present data in both Polish

and Russian contexts are ‘koroče’ – lit. ‘shorter’ – *in short*, ‘tipa’ – *sort of* and ‘prosto’ – *simply*. Just like ‘nu’ and ‘vot’, these three discourse markers are characterised by regular and frequent distribution across the corpus, although ‘prosto’ is a bit less common than ‘koroče’ and ‘tipa’. To put it in quantitative terms, ‘koroče’ is employed on 306 occasions by nineteen speakers, ‘tipa’ is equally popular with a total of 364 instances of use across all the speakers in the data, while ‘prosto’ is used on 72 occasions by seventeen different speakers.

While ‘koroče’ and ‘tipa’ feature more prominently in the corpus than ‘prosto’, it seems plausible to test all three of them against the hypothesis of ‘fusion’. It can be formulated as follows: if Russian discourse markers ‘koroče’, ‘tipa’ and ‘prosto’ can appear as single switches in a Polish base, and have no competing equivalents from Polish, then they can be regarded as discourse markers that have been ‘fused’ into the emerging Polish-Russian mixed code on a par with discourse particles ‘nu’ and ‘vot’. Should this hypothesis prove to be correct, it could be concluded that the Polish-Russian language alternation reflected in the present data fulfils one of the criteria to be classified as an emerging mixed code, while the three discourse markers, along with the two discourse particles could be classified as lexical items that have undergone an even further development and became part of the fused lect.

In order to test this hypothesis further, the following questions will be addressed in the data analysis: (1) Are there any Polish equivalents that have they been attested in the present data? (2) What functions do these discourse markers perform? (3) Are they able to appear as stand-alone items (i.e. single switches) in a Polish base?

4.5.1 ‘Koroče’

With a total number of 306 instances, ‘koroče’ is one of the most popular discourse markers in the present corpus. According to Gorbachevich’s (2004) dictionary, it is an adverb in its comparative degree, derived from ‘korotko’ - *shortly, briefly*.

When employed with the verb ‘govorja’, it functions as ‘vvodnoe slovosocetanie’ – *introductory collocation*, i.e. a Russian equivalent of a discourse marker, and is employed to signal the speaker’s rhetorical stance towards what is being said. As far as the present corpus is concerned, ‘koroče’ always appears on its own without the verb, but it never functions as an adverb. A closer look at the data indicates that ‘koroče’ functions metalingually, and expresses the same meaning as the extended collocation ‘koroče govorja’. Therefore, it is plausible to assume that this discourse marker started as ‘koroče govorja’ but as a result of frequent usage became grammaticalised into the shorter discourse marker ‘koroče’.

The closest equivalents from Polish are ‘krótko mówiąc’ – *briefly speaking*, ‘jednym słowem’ – *in one word* and ‘w sumie’ – *in sum*. However, neither of the three is employed in the corpus. This suggests that ‘koroče’ is not only the default and the sole option that the speakers resort to, but also constitutes a strong initial piece of evidence for the hypothesis of ‘fusion’.

Additional evidence in favour of ‘fusion’ comes from the fact that in 196 instances out of a total of 306 ‘koroče’ can appear as a single switch in the Polish base. Moreover, it does not appear to have any syntactic restrictions and can be used in sentence initial, medial and final positions. ‘Koroče’ in the sentence initial and final positions is exemplified in Extract 15, while in sentence medial position in Extract 16:

Extract 15

- (1) D: _[RUS] **me-me- mestna** ryżska taka łódeczka
 lo-lo- local from Riga such little boat
 ‘lo- lo- local little boat from Riga’

- (2) C: aaa
 aaa
 ‘aaa’

- (3) D: _[RUS] **koroče** nas nie chciał potem wie—
in short us not wanted then ta—
‘in short [he] then didn’t want to ta[ke] us’
- (4) C: wszędzie woda
everywhere water
‘water everywhere’
- (5) D: mhm (.) nie chciał wieźci na lotnisko nas _[RUS] **koroče**
mhm (.) not wanted take on airport us in short
‘mhm (.) [he] didn’t want to take [us] to the airport in short’

In the above extract the sentence initial ‘koroče’ appears in turn (3), while the sentence final in turn (5). In both instances it occurs in an otherwise Polish base. As far as its function is concerned, ‘koroče’ in turn (3) can be described as introducing a new topic. The speaker wishes to tell the story about his adventures in Riga, and he therefore foregrounds this new narrative with the help of a discourse marker. However, as soon as he begins his story, he gets interrupted by speaker C. Having briefly acknowledged C’s comment, D resumes his story in turn (5), where the second instance of ‘koroče’ can be found. It seems to perform a two-fold function here. On the one hand, it signals the resumption of the previous topic of conversation following a brief digression initiated by the listener. On the other hand, it functions as a turn-closing device, thus helping the speaker to segment his narrative into smaller parts.

Another example can be found in Extract 16:

Extract 16

- (1) E: to ten Skansen _[RUS] **tipa** jak on (0.4) mhhh (.) _[RUS] **nu** tam
it that Skansen type how he (0.4) mhhh (.) well there
‘it’s this Skansen, what it’s called (0.4) well there is’
- _[RUS] **koroče** i park zoologiczny i— _[RUS] **nu koroče vsja** jakby licz
in short and park zoological and— well in short all as if count
[LT] Švedija
Sweden

‘in short a zoological park and— well in short the whole of Sweden.’

In the above extract there are two instances of ‘koroče’, which occur in a sentence-medial position. In the first instance ‘koroče’ appears as a single switch in the Polish base. The function that it performs can be described as ‘summarising’ and is most closely related to the literal meaning of ‘koroče’ – *in short*. The speaker starts explaining what Skansen is and therefore attempts to list all of its uses under a brief description, which he precedes with ‘koroče’. However, after the first option the list gets interrupted. Instead, the speaker formulates a different summary, which is preceded with the discourse particle ‘nu’, followed by the discourse marker ‘koroče’. In the second instance ‘koroče’ appears as part of a longer intraclausal switch from Russian.

While in the above extracts it is possible to identify a tentative function performed by different instances of ‘koroče’, it should be noted that it is not always apparent. There are instances in the data where this lexical item seems to function only as a discourse-structuring device, and can be likened to a spoken punctuation mark. An example of this type of usage can be found in Extract 17:

Extract 17

- (1) M: _[RUS] **koroče** on tym _[RUS] **trehtonnikom** chciał ich (.) _[RUS]
in short he this three-tone wanted them (.)
povydergivat’ (.)
pull out
‘in short he wanted to pull [them] out with this three-tone vehicle (.)’
może gdzie wtedy i powyłaziło **vse** **koroče** guz na tym **vo**
maybe where then and came out everything in short bump on this well
(.) na kole (0.2)
(.) on wheel (0.2)
‘maybe [that’s when] it all came out, in short a bump on this well (.) wheel’
_[RUS] **koroče** (.) _[RUS] **polučilos’** tak że nowa _[RUS] **rezina na pogruzčik**
in short (.) turned out so that new rubber on loader

kupić
to buy

‘in short it turned out that to buy the new rubber for the loader’

[RUS] **koroče** (0.1) kosztuji (0.2) [RUS] **trista** **sem’desjat** litów plus
in short (0.1) costs (0.2) three hundred seventy litas plus
[LT] PVM (.)
VAT

‘in short (0.1) it costs (0.2) three hundred and seventy litas plus VAT’

[RUS] **koroče** wyszło [RUS] **pjat’ soten** (0.1)
in short turned out five hundreds (0.1)

‘in short, it turned out five hundred’

[RUS] **koroče** (.) mówi (.) [RUS] **nu ja vam daju vremja poumat’**
in short (.) says (.) well I you give time to think

‘in short, he says (.) well, I give you some time to think.’

The above extract is a good example of what seems to be the emerging Polish-Russian mixed code, for it contains numerous and frequent switches from Russian inserted into the Polish base. The discourse marker ‘koroče’ is employed on six occasions. In the majority of cases it appears at conversational action boundaries, i.e. beginning of a new topic (line (1)) and transition between utterances within a turn (lines 3, 4, 5 and 6), accompanied by pauses either on both, or only on one side. However, this example is significant not only because it is representative of the Polish-Russian mixed code, but also because it illustrates priming and its influence on quantitative analysis. Priming refers to the short-lived memory of lexical and structural units that is encountered in language production (Bock 1986). Producing a sentence involves the activation of procedures associated with production of a particular syntactic form. This activation does not disappear immediately, thus speakers tend to repeat the same syntactic structures since these forms remain activated and therefore easily accessed within a certain period of time. This appears to be the case in the example above, where each sentence within a turn begins in exactly the same way, i.e. it is preceded by the same discourse marker. Since this study is based on the corpus-driven, rather

than production-driven approach, each instance of ‘koroče’ employed in the above example was quantified as a separate token. However, from the purely psycholinguistic perspective, it could be argued that the recurrent use of the Russian discourse marker is the result of priming, which would also entail a different quantification.

The above extract also demonstrates that ‘koroče’ can occur in a cluster with Russian lexical items (line (2)), as well as as a single switch in the Polish base (lines (1), (4), (5)). Those instances where the discourse marker is singled out by pauses on both sides, namely in lines (3) and (6) are a bit less straightforward, for it can be argued that the use of ‘koroče’ triggers further switching (line (3)), or that the presence of other Russian lexical elements calls for the employment of a Russian discourse marker (line (6)). However, the appearance of pauses is significant here, for they separate the discourse marker from both, the preceding and following discourse and it therefore seems plausible to treat it as a single switch. This line of reasoning is further reinforced by the fact that in line (3) ‘koroče’ is preceded by Polish discourse and followed by Russian discourse, whereas in line (6) it is preceded by Russian discourse and followed by Polish discourse, which demonstrates that there is no apparent pattern that could be singled out here.

As far as ‘koroče’ is concerned, the examples suggest that it has been ‘fused’ into the Polish-Russian mixed code on a par with the discourse particles ‘nu’ and ‘vot’. This finding is further corroborated by the quantitative evidence, namely the fact that it is employed on 306 occasions by nineteen speakers. Moreover, there are no equivalents from Polish, which suggests that ‘koroče’ does not have any competition and is the default option for all the speakers in the data. Finally, it can appear as a single switch in the Polish base, which means that it has been accepted as part of a switched mode of speaking, prevalent in the Polish speech community investigated in the present thesis.

4.5.2 ‘Tipa’

With a total number of 364 instances, ‘tipa’ – *sort of* is another popular discourse marker that can be found in the present corpus and an equally good candidate for the hypothesis of ‘fusion’. It is derived from the noun ‘tip’ – *type*, hence the literal translation of ‘tipa’ would be *of the type*. The closest equivalent in English is the discourse marker *sort of*, which has been analysed in great detail by Holmes (1988) and Aijmer (2002: 175–209). Holmes (1988: 94) introduces a distinction between the evidential (epistemic modal) and affective (interpersonal) function of *sort of*. In the evidential function *sort of* is used when the speaker wishes to avoid being precise. This may occur because he or she cannot find the right word or because there is a lexical gap in their mental lexicon. On the other hand, the affective *sort of* is hearer-oriented and closely associated with claiming common ground and avoiding disagreement. Aijmer (2002: 209) further adds that “the affective *sort of* has softening and polite functions; it contributes to the informality of conversation and creates a congenial atmosphere”.

It seems that the two-fold functional division proposed for the English *sort of* can also be applied to the Russian ‘tipa’. Levontina and Shmelev (2007: 260) categorise it as “a marker of rough-and-ready nomination or of evading responsibility in case the statement lacks precision”. They further explain that by employing a discourse marker of this type the speaker admits “the possible inaccuracy of a nomination and refuses responsibility for it” (2007: 263). They also briefly mention the affective function of ‘tipa’ and report that by employing this discourse marker speakers create informality in their conversational exchanges (2007: 265).

The closest equivalent from Polish is ‘typu’ - lit. ‘of the type’ – *sort of*, derived from the noun ‘typ’ – *type*. However, unlike the Russian ‘tipa’, it seems that the Polish ‘typu’ has not been grammaticalised into a discourse marker, which is corroborated by a search through the entries of the National Corpus of Polish. In the examples

of spoken Polish that can be found in the corpus ‘typu’ is always found in its literal meaning, e.g.: ‘zapisy tego typu’ – *records of this type*, ‘tego typu rozwiązania’ – *solutions of this type*. Interestingly enough, the Polish ‘typu’ does appear in five instances in the present data. However, it is used as a discourse marker, and therefore in analogy with the Russian ‘tipa’. An example of such usage is discussed in greater detail in Extract 21.

On the basis of the quantitative distribution, regularity of occurrence and lack of a direct functional equivalent from Polish it is possible to consider ‘tipa’ as a good candidate for the hypothesis of ‘fusion’. Moreover, this discourse marker also meets another criteria for ‘fusion’, namely the fact that it can occur as a single switch in the Polish base (272 instances out of a total of 364). Similarly to other Russian discourse markers, ‘tipa’ can be found in utterance initial, medial and final positions, which attests to its syntactic freedom and integration within the host language.

As stated earlier, the functions of ‘tipa’ can be likened to those of the English discourse marker *sort of* and can be roughly divided into two types, the evidential and the affective (Holmes 1988: 94). Examples of the evidential ‘tipa’ can be found in Extracts 18, 19 and 20 below:

Extract 18

- (1) C: zaraz zobaczysz ^[RUS] **tut** taki żołnierz ^[RUS] **tipa** (.) który tam ^[RUS]
soon will see here such soldier sort of (.) which there
ohranjaet
protect
‘soon you will see here sort of a soldier (.) who protects [the place] there’

Extract 19

- (1) D: a tutaj wiesz ^[RUS] **tipa** amerykańskie ^[RUS] **goročki** takie
and here know sort of american hills such
‘and here, you know, sort of a roller-coaster’

In both examples the discourse marker ‘tipa’ appears in sentence-medial positions as a single switch. It does not trigger further switching, and in both instances it is inserted into the Polish base. In Aijmer’s terminology, the Russian ‘tipa’ that can be found in the above examples can be described as hedging ‘lexical imprecision’ (2002: 197), which usually occurs when one lacks the vocabulary needed to talk on a particular topic. With the help of ‘tipa’, the speaker conveys that his description is vague and not as precise as it could have been. Thus, the term ‘żołnierz’ – *soldier* is a close description, but not exact, for the speaker probably meant to say *guard*. Likewise, ‘amerykańskie [RUS] goročki’ is only an approximate description of a roller-coaster that the speaker saw.

A few more examples of the evidential ‘tipa’ can be found in Extract 20:

Extract 20

- (1) D: _[RUS] **tipa** tam pisało się że (.) żeby nie wychodzić na ulica _[RUS]
 sort of there wrote itself that (.) that not leave on street
tipa oni w środku
 sort of they in inside
 ‘it was sort of written there (.) not to go outside sort of [because] inside they’
 wykopywali _[RUS] **tipa** jeżeli kto na ciebie napadałby czy coś
 dug out sort of if someone on you attacked or something
 (.) żeby miałby ty wody _[RUS] **tipa**
 (.) that had you water sort of
 ‘were digging out sort of if someone attacked you or something (.) so that you
 had some water sort of’

In the above extract ‘tipa’ appears as a single switch in an otherwise Polish base in all three positions, initial, medial and final, which further corroborates the view that it can be regarded as a ‘fused’ lexical item. In the above extract speaker D is trying to recount the content of a health and safety note he saw in a museum. The extract contains four occurrences of ‘tipa’, all of which function as hedges to indicate that the information provided by D may not be precise.

The function of ‘tipa’ exemplified above is also characteristic of the five instances of the Polish ‘typu’, which can be found in the present corpus. They feature in the discourse of two speakers, J and K. In all instances the Polish ‘typu’ seems to be employed in the same way as the Russian ‘tipa’ in its evidential function, which suggests that the use of the Polish ‘typu’ has been extended on the model of its Russian equivalent, and that it can therefore be regarded as a case of semantic extension of the Russian particle. One such example can be found in Extract 21:

Extract 21

(1) J: u mniej jedna koleżanka _[RUS] **vydvinuła** taka teoria=
 at me one friend pushed such theory=
 ‘one of my friends proposed such a theory=’

K: _[RUS] **nu**
 O.K.
 ‘O.K.’

J: =typu (.) że ludzie sami winni w tym że jest zimno
 =sort of (.) that people themselves guilty in this that is cold
 ‘=sort of (.) that people themselves are to blame for feeling cold’

Speaker J employs ‘typu’ as a discourse marker in order to signal to the interlocutor that the information provided may not be accurate and precise. As stated earlier, this type of usage is not found in Standard Polish, which suggests that the Polish ‘typu’ can be classified as a semantic extension of the Russian particle.

Finally, Extracts 22 and 23 contain examples of the affective function of ‘tipa’. Aijmer (2002: 202) explains that the frequent use of *sort of* can signal a close relationship between speakers, for “it is linked to the hearer and to affect by means of copresence or common ground”. As James (1983: 202) puts it, such discourse markers as *sort of* and *like* are sometimes employed to establish ‘social synonymy’ between interlocutors. This function of *sort of* can also be attributed to those instances in the data where ‘tipa’ does not perform any other immediately identifiable functions. Some examples

can be found in Extracts 22 and 23:

Extract 22

D explains why he didn't attempt to stroke any animals in the zoo.

- (1) D: nie bardzo tam _[RUS] **tipa** można (.) tam _[RUS] **tipa** chodzo patrzajo
not very there sort of can (.) there sort of walk look
tak trochi wiesz
like this a bit know
'you can't really (.) [they] sort of walk there and watch a bit, you know'

Extract 23

M talks about the request made by his employer.

- (1) M: najpierw _[RUS] **tipa** odświecił sobie ta _[RUS] **bumaga** (0.1)
first sort of photocopied for himself this paper (0.1)
_[RUS] **koroče** (.) od— (0.1)
in short (.) from— (0.1)
'first he photocopied that piece of paper for himself (0.1) in short (.) from—'

z drugiej godziny _[RUS] **tipa** napisać _[RUS] **objasnitel'noe** i pół do
from second hour sort of write explanatory and half to
trzeciej jemu przywieźć
three him bring
'from two o'clock sort of write an explanation and bring it to him at half past
two'

In the two extracts above the Russian 'tipa' appears as a single switch in the Polish base, which supports the hypothesis of 'fusion'. It does not appear to perform any evidential function whereby the speaker wishes to signal lexical or propositional imprecision, vagueness or uncertainty. Neither does it structure the speakers' turns in a way that 'koroče' does in Extract 17. The only function that seems to be relevant here is that of establishing 'common ground' with the interlocutor and maintaining an informal character of the conversation.

As far as ‘tipa’ is concerned, it can be concluded that it has been fused into the Polish-Russian mixed code on a par with ‘koroče’, ‘nu’ and ‘vot’. This is corroborated by the quantitative evidence and its equal distribution across all the speakers in the corpus. Moreover, it has no immediate functional equivalents from Polish, which suggests that ‘tipa’ is the default option that the speakers resort to. Finally, it can appear as a single switch in the Polish base, which means that it is regarded as part of what appears to be the Polish-Russian mixed code.

4.5.3 ‘Prosto’

As far as its quantitative distribution is concerned, the discourse marker ‘prosto’ — *simply* is employed on 72 occasions by seventeen speakers. According to Gorbachevich’s (2004) dictionary, it can function as an adverb with a range of meanings, namely *simply, without any difficulty or effort; frankly, openly; incidentally, without special intention*, as well as as a discourse marker to reinforce either the meaning of the lexical item that it modifies, or the meaning of the entire proposition. Interestingly enough, in the present corpus ‘prosto’ is employed only as a discourse marker.

The Russian ‘prosto’ does have a Standard Polish cognate ‘prosto’, which only functions as an adverb and expresses such meanings as *straight; upright; ordinary and uncomplicated; without delay*. However, it can also be employed as a discourse marker, but in a slightly altered form, i.e. as a propositional phrase ‘po prostu’. The Polish adverb ‘prosto’ does not feature in the present corpus, but the discourse marker ‘po prostu’ does, although its distribution is more limited and idiolectal compared to the Russian ‘prosto’. While ‘po prostu’ is employed on fifteen instances by three different speakers, namely J, K and L, as many as thirteen instances belong to speaker J. It appears that this speaker is the only speaker in the corpus who displays a clear preference for the Polish discourse markers, although she also does use the Russian ‘prosto’ on one occasion. Speakers K and L, on the other hand, seem to favour the

Russian variant (four occurrences for speaker L and eight occurrences for speaker K) and employ the Polish ‘po prostu’ on one occasion each.

Although the above observations show that in the present data the Polish and Russian discourse markers ‘prosto’ and ‘po prostu’ co-exist, the latter is clearly idiolectal, which suggests that the former can still be tested for the hypothesis of ‘fusion’ on a par with ‘tipa’, ‘koroče’, ‘nu’ and ‘vot’. Thus, if the discourse marker ‘prosto’ is employed on a frequent and regular basis by all, or at least most, speakers in the data without any significant competition from Polish, and if it can appear in the Polish base without any syntactic restrictions, then it can be classified as another discourse marker that has been ‘fused’ into the emerging Polish-Russian mixed code.

Just like the other four discourse markers analysed above, ‘prosto’ can be inserted into the Polish base (a total of 51 instances), as well as be employed in the Russian discourse (a total of 21 instances). When this discourse marker is used in the Polish base, it does not appear to have any syntactic restrictions and can be found in sentence initial, medial and final positions. Some examples can be found in Extracts 24, 25 and 26:

Extract 24

- (1) F: _[RUS] **v osnovnom** takie ludzi _[RUS] **postarše** pracuju
mainly such people older work
‘mainly such people [who are] older, they work’
- (2) G: _[RUS] **prosto** gada jeszcze zechciał profesja jedna _[RUS] **dostat’** **govorit**
simply says more wanted profession one get says
‘simply [he] says [that] he wanted to get one more profession, [he] says’

From the semantic and syntactic point of view, the discourse marker in the above example appears to be redundant, for it does not add anything to the propositional content of the utterance. It is difficult to pin-point the exact pragmatic function that

it performs, as it merely seems to reinforce the overall proposition expressed by the speaker. Moreover, the use of the Russian discourse marker in the sentence-initial position does not trigger further switching, and the rest of the turn continues with Polish as the base language, with two single switches from Russian inserted at the very end. This shows that the choice of a Russian discourse marker is non-consequential on the level of language choice.

An example of ‘prosto’ in sentence-medial position can be found in Extract 25:

Extract 25

- (1) K: _[RUS] **nu** (.) bardzo duzo wszystkiego (.) jednocześnie
 well (.) very much everything (.) simultaneously
 ‘well (.) a lot of everything (.) simultaneously’
- (2) J: tak (.) i wtedy nic nie masz (.) to musisz _[RUS] **prosto** sobie
 yes (.) and then nothing not have (.) so must simply yourself
 wypisać priorytety
 write down priorities
 ‘yes (.) and then you don’t have anything (.) so you must simply write down
 priorities for yourself’

In the above extract the discourse marker ‘prosto’ appears in turn (2) as a single switch in an otherwise Polish base. It is difficult to identify the exact discourse-related function here but it seems that, just like in the previous example, it is employed in order to reinforce the overall proposition expressed.

Finally, an example of ‘prosto’ in the sentence-final position can be found in Extract 26:

Extract 26

W is taking about car repair services

- (1) W: trzeba takich mieć znajomych które remontują (.) żeby ty mógł
 must such have acquaintances who repair (.) so that you could

‘[one] must have such friends who do repairs (.) so that you could’
 zapędzić i żeby oni tobie zrobili (.) trzeba szukać ^[RUS] **prosto**
 drive and so that they you did (.) must search simply
 ‘drive [you car there] and so that they would do it for you (.) one must simply
 search’

In the above example the Russian discourse marker is employed as an utterance-closing device, which performs an emphatic function in order to reinforce the overall meaning of the proposition. Moreover, the lexical item ‘prosto’ could be omitted without changing the meaning of the sentence and rendering it ungrammatical.

According to the discussion presented above, the Russian discourse particles ‘koroče’ – lit. ‘shorter’ – *in short*, ‘tipa’ – *sort of* and ‘prosto’ – *simply* are equally good candidates for the hypothesis of ‘fusion’ as the discourse particles ‘nu’ and ‘vot’. Apart from ‘prosto’, the discourse markers analysed in this section do not have any immediate competing forms from Polish. While the Polish cognate of ‘prosto’ does appear in the present corpus, it features predominately in the discourse of one speaker, which suggests that it is highly idiolectal. This does not contradict the hypothesis of ‘fusion’, as for all the remaining speakers it is the Russian discourse marker that constitutes the default option. Just like ‘nu’ and ‘vot’, the three discourse markers can appear in either the initial, medial or final position in a sentence. Moreover, they can be used as stand-alone items in the Polish base, which means that speakers have accepted them as part of their lexicon. These observations can be treated as evidence of the hypothesis of ‘fusion’ in what Matras refers to as its ‘diachronic manifestation’, though a bit less so for ‘prosto’, where ‘po prostu’ is still employed by some speakers. Since Russian swearwords ‘blin’ – *damn* and ‘bljad’ – *whore* tend to act as discourse-structuring devices, and thus perform similar functions to those of discourse markers, it seems plausible to extend this hypothesis even further and include these two forms as well. They will be analysed in greater detail in the following subsection.

4.6 Swearwords ‘blin’ and ‘bljad’

According to Andersen and Trudgill (2007: 195), swearing can be defined as language use in which the lexical item or expression refers to something taboo or stigmatised, expresses strong emotions or attitudes and is not to be interpreted literally. As demonstrated in the existing studies, just like other forms of linguistic practice, swearing can be described as very versatile, for it can fulfil a range of functions, such as expressing emotion, creating humour and verbal emphasis, achieving social bonding and solidarity, as well as constructing and projecting a desired identity (Stapleton 2010: 290). In addition to interpersonal, social and psychological functions, swearwords can also perform pragmatic or discourse-related functions. According to Dewaele (2004: 205), they can be described as “multifunctional, pragmatic units which assume, in addition to the expression of emotional attitudes, various discourse functions. They contribute, for instance, to the coordination of the interlocutors, the organisation of the interaction, and the structuring of verbal exchange; in that, they are similar to discourse markers.”

The similarity between swearwords and discourse markers is conditioned by the fact that lexical items attributed to both of these categories arise as a result of the same process, commonly referred to as ‘grammaticalisation’. Following Hopper and Traugott (2003), grammaticalisation occurs when content words and larger clausal structures lose their meaning and develop into function words. As far as discourse markers are concerned, probably the most illustrative example discussed in the present data is the grammaticalisation of the extended collocation ‘koroče govorja’ – *briefly speaking* into the discourse marker ‘koroče’ – *in short*. In the case of swearwords, the process of grammaticalisation is further accompanied by two other processes, known as ‘deategorialisation’ and ‘desemanticisation’ (Ljung 2011: 21). Desemanticisation is characterised by the loss of meaning in such words as *hell* or *shit* when they are used in swearing. The same processes also applies to such religious terms as *God*

and *Christ* when they feature as interjections *Oh God!*, *Christ!*. As Ljung (2011: 21) explains, while these lexical items can still be used in their literal meaning, they are now far outnumbered by the interjectional uses, in which sacred names are employed to convey the speaker's reaction to some sort of stimulus and function as pragmatic markers which express the speaker's stance towards the proposition. He concludes that "the processes of decategorialisation and desemanticisation work in tandem in the development leading to swearing", for these lexical items, which started as proper nouns, changed their category membership to interjections at the same time as the nature of the religious terms God and Christ became desemanticised.

The same processes apply to such terms as the Russian 'bljad'. While it can still be employed with the meaning of *whore* as a pejorative person-denoting noun, due to the frequency of use, its literal meaning has become desemanticised. It can be further argued that 'bljad' was also subject to the processes of decategorialisation, whereby its interjectional usage started to prevail over its nominal usage. This can easily be confirmed by running a search in the spoken component of the Russian National Corpus. The sub-corpus of spoken Russian contains a total of 232 instances of 'bljad', which are made up of 225 interjections and only seven nouns. According to Ljung (2011: 43–44), in such interjectional speech acts as the Russian 'bljad!', the Polish 'kurwa!' or the French 'putain!' there is no reference to either prostitutes or whores, but these lexical items are used merely as "vehicles for the speaker's feelings". He therefore labels them as 'expletive interjections', which are considered to be cathartic, i.e. not aimed at others, but serving as "outlets for the speaker's reactions to different mishaps and disappointments" (Ljung 2011: 30). However, they also have many non-cathartic uses and are widely used to carry out various pragmatic and discourse-related functions, such as acknowledging the information given in a previous turn, emphasising the relevance of the propositional content, or expressing the speaker's attitude or epistemic stance (Ljung 2011: 87–91).

The discourse functions of ‘blin’ and ‘bljad’ are briefly addressed by Levontina and Shmelev (2007: 8), who propose the term ‘bljakanje’ to describe the linguistic practice of the frequent insertion of the two swearwords. They claim that ‘bljakanje’ helps to maintain the rhythmic structure of speech and expresses the general attitude of the speaker. While ‘bljad’ is considered vulgar and taboo, ‘blin’ can be employed as its softer replacement. According to McArthur (1992: 661), these replacements can be of two kinds: 1) replacements that create a nonsense equivalent of a swearword, e.g. *Gosh* for *God* or *Gee* for *Jesus*; 2) replacements that substitute an everyday expression of similar sound and length, e.g. *darn* instead of *damn* or *shoot* instead of *shit*. However, as Ljung (2011: 11) correctly points out, strictly speaking, it is misleading to treat such euphemistic replacements in terms of creating an equivalent of a swearword. He explains that “what *Gosh!* and *Gee!* replace is not the words *God* and *Jesus*, but interjectional utterances containing these words which are used in swearing. Outside of the interjectional context, *God* and *Jesus* may never be replaced by *Gosh!* and *Gee!* [...]” Thus, ‘blin’ can be described as the interjectional euphemism for ‘bljad’, which was presumably chosen for its similarity in length and sound.

Since swearwords, which occur as expletive interjections tend to perform discourse-related functions, it seems reasonable to test them against the hypothesis of ‘fusion’ along with discourse markers and discourse particles. ‘Blin’ and ‘bljad’ can be singled out as particularly strong candidates for they fulfil all of the required criteria. First of all, both of them can occur as single switches in the Polish base. Secondly, they feature in the present corpus with great frequency and regularity. The Russian ‘bljad’ is employed on 692 occasions by seventeen speakers, whereas its milder euphemism ‘blin’ is used on 128 occasions by fourteen speakers. Third, the two swearwords are used exclusively without any competition from Polish. While there is one instance of the Polish equivalent ‘kurwa’ – *whore* in speaker A’s discourse, it is used as a pejorative noun with its literal meaning, and not as an expletive interjection. In fact,

the Russian ‘blin’ and ‘bljad’ are only ever employed as expletive interjections in the present data, which suggests that they perform discourse-related functions and further emphasises their similarity to discourse markers and discourse particles.

Like discourse markers and particles analysed in previous sections, the swearwords ‘blin’ and ‘bljad’ can appear in the Polish base without any syntactic restrictions. Both of them can occur in sentence initial, medial and final positions. Some examples can be found in Extracts 27–30:

Extract 27

- (1) W: ja pracuja zawsze do dziewiętej (.) jedenastej (.) u mnie nie ma
 I work always until nine (.) eleven (.) at me no is
 takiego żeb byłoby do szóstej (.) piątej
 such so that would be until six (.) five
 ‘I always work until nine (.) eleven (.) [I] never happen to work until six (.)
 five’
- (2) X: _[RUS] **blin** ale też źle (.) jeżeli na praca trzeba to do szkoły już
 damn but also bad (.) if to work must then to school already
 nie idzisz
 not go
 ‘damn, but [this is] also bad (.) if [you] must go to work then you don’t go to
 school’

In the above extract ‘blin’ occurs in the sentence-initial position. The Russian lexical item does not trigger further switching as the base of the utterance remains Polish. When used in sentence-initial positions, expletive interjections have scope over the proposition underlying the clause, and they tend to express subjectivity, the speaker’s attitude and/or epistemic stance (Ljung 2011: 88). This observation seems to apply in the above example, where the Russian ‘blin’ is employed just before the expression of the speaker’s subjective evaluation of the information provided in the previous turn.

Another example can be found in Extract 28:

Extract 28

- (1) A: jak można na tyle _[RUS] **bljad'** pokłócić się żeby miesiąc z
how can on so much whore argue so that month with
sobo nie rozmawiać
each other not talk
'how can you, whore, argue to such an extent that you don't talk to each other
for a month?'

In the above example the swearword 'bljad'' is employed as an expletive interjection in the sentence-medial position, where it occurs as the only lexical item from Russian in an otherwise Polish base. The expletive interjection has scope over the verb 'pokłócić się', which suggests that its primary discourse function can be described as emphatic, for the swearword adds additional emphasis to the following verb.

Another example of 'bljad'' in the utterance-medial position can be found in Extract 29:

Extract 29

- (1) Z: dlaczego on nie chciał?
why he not wanted
'why he didn't want?'
- (2) Y: ni wiem (.) _[RUS] **vot** (.) i::: (.) _[RUS] **bljad'** (.) co ja chciała
not know (.) there (.) and (.) whore (.) what I wanted to say
'Don't know (.) there (.) a:::nd (.) whore (.) what did I want'
- powiedzieć (.) aj (.) i ona _[RUS] **v Bel'gii** skończyła _[RUS] **čego-to** i
to say (.) oh (.) and she in Belgium finished something
'to say (.) oh (.) and she finished something in Belgium'
- i popadła gdzieśkim kimścik tam pracować
and got somewhere someone there to work
'and got [a job] somewhere as someone'

The Russian swearword can be described as a separate utterance within speaker Y's turn, which is clearly set off by pauses on both sides from the preceding and following discourse. In the existing research on discourse analysis the expletive interjection 'bljad'', as exemplified above, would normally be described as a 'gap filler', which helps the speaker win some time and to recall what he was going to say. This observation is further corroborated by the presence of the discourse particle 'vot', which is employed to carry out a very similar function. As far as the hypothesis of 'fusion' is concerned, the Russian 'bljad'' constitutes a separate utterance within a longer turn. However, it does not trigger further switching, as discourse that follows has Polish as its base language.

Some further examples can be found in Extracts 30:

Extract 30

- (1) C: temu dziecku ^[RUS] **bljad'** (.) jemu dziesięć lat (.) a u niego oczy
 this child whore (.) he ten years (.) but at he eyes
 człowieka
 man
 'this child (.) whore (.) he's ten years old (.) but [he's got] eyes of a man'
 który ^[RUS] **bljad'** pół życia przeżył
 who whore half life lived
 'who, whore, has [already] lived half of his life'

The above extract contains two instances of 'bljad'', both of them inserted in the turn-medial positions into the Polish base. Speaker C's turn begins with the collocation 'temu dziecku', which appears on its own in the utterance-initial position, where it is followed by the swearword 'bljad'' and a pause. The second utterance starts right after the pause and refers back to the child. Such a syntactic construction, with the object singled out in the beginning and reiterated by the pronoun, creates additional emphasis, which is further reinforced by the presence of the swearword. While in Extract 28 the expletive interjection 'bljad'' occurs in the beginning and has scope

over the entire proposition, in the above example it appears to have a retroactive scope over the preceding phrase. Its function can therefore be described as ‘emphatic’ and discourse-structuring. In the second instance ‘bljad’ is employed in the clause-medial position where it has scope over the remaining part of the utterance and adds further emphasis.

In addition to discourse-related functions, the presence of the swearword seems to also indicate the speaker’s feelings and emotions. However, expletive interjections may be employed to convey a range of feelings, such as irritation, anger, surprise, or joy, to mention just a few. As Ljung (2011: 86) explains, expletive interjections can be prompted by a mixture of emotions, but it is impossible for an outside observer to determine which of these feelings has the upper hand in each specific case.

Some examples of ‘bljad’ used in a sentence-final position can be found in Extract 31:

Extract 31

Speaker A is talking ironically about her husband.

- (1) A: już piąta (.) ja po pracy ^[RUS] **bljad’**
 already five (.) I after work whore
 ‘it’s already five (.) I am after work, whore’
- (2) B: mąż czeka?
 husband waits
 ‘is your husband waiting?’
- (3) A: mąż śpi ^[RUS] **bljad’** (.) u mnież mąż ^[RUS] **bljad’ ofigennyj**
 husband sleeps whore (.) at me husband whore awesome
robotnik
 worker
 ‘husband is sleeping, whore (.) my husband, whore, is an awesome worker’
- (4) B: jaki? ^[RUS] **ofigennyj**?
 what awesome

‘what [worker]? awesome’

- (5) A: _[RUS] **da** (.) **ofigennyj** (.) **trudogolik** **bljad**
yes (.) awesome (.) workaholic whore
‘yes (.) awesome (.) a workaholic, whore’

In the above extract there are four instances of ‘bljad’’: two in the turn-final position (turns (1) and (5)), one in the utterance-final position within a turn (turn (3)) and one in utterance-medial position (turn (3)). On the one hand, regular presence of swearwords within a relatively short conversational exchange reflects the speaker’s emotions, which could possibly include irritation, anger or annoyance. On the other hand, frequent insertions of ‘bljad’ help the speaker create structure within her discourse and add emphasis. The sentence-final instances of ‘bljad’ at the end of turn (1) and within turn (3) can be likened to a spoken punctuation mark, which marks the end of a sequence. The second instance of ‘bljad’ in turn (3), as well as the sentence-final ‘bljad’ in turn (5) appear to perform an emphatic function. In turn (3) ‘bljad’ has scope over ‘ofigennyj rabotnik’, whereas in turn (5) it has retroactive scope over the preceding word ‘trudogolik’, which means that the presence of the swearword provides additional emphasis. Moreover, the above extract also demonstrates that ‘bljad’ can occur in a cluster with Russian lexical items (turns (3) and (5)), as well as as a single switch in the Polish base (turn (1) and (3)). This shows that its presence is equally accepted within Russian, as well as Polish discourse.

The examples discussed above demonstrate that, in addition to discourse particles and discourse markers, the swearwords ‘blin’ and ‘bljad’ are two other lexical items that can be classified as good candidates for the hypothesis of ‘fusion’. Both swearwords appear in the present corpus with great frequency and regularity without any immediate competing forms from Polish, which suggests that they constitute the default option. Moreover, they can appear without any syntactic restrictions in the initial, medial or final position of a sentence. Finally, they can be used as stand-alone items

in the Polish base, which means that the speakers have accepted them as part of their lexicon. These observations can be treated as evidence that ‘blin’ and ‘bljad’ have been ‘fused’ into the Polish-Russian mixed code alongside Russian discourse markers and particles.

4.7 Conclusions

The extracts presented in this section show that Russian discourse markers, which also include discourse particles and expletive interjections make up one group of lexical items that are frequently inserted in what can be qualified as the emerging Polish-Russian mixed code. According to the first premise of the Emerging Mixed Code/Fusion Hypothesis, if Russian discourse markers can be inserted into the Polish base without triggering further switching, have no competing equivalents from Polish and are exclusively employed in both Polish and Russian contexts, then this can be taken as one piece of evidence in favour for the hypothesis tested in this thesis. On the basis of examples analysed above it can be concluded that the premise tested in this chapter is borne out by the Polish-Russian data.

One of the elementary, but crucial findings of the analysis is that Russian discourse markers appear regularly as single switches in an otherwise Polish base without triggering further switching. This in turn demonstrates that these elements have been sufficiently integrated to be inserted into utterances that are otherwise monolingual. This finding is consistent with the hypothesis tested in this chapter, as well as the prevailing assumption that is dominant in the existing literature as to the special status of discourse markers. Due to their syntactic and semantic detachment from the propositional content of the utterance, discourse markers have often been placed at the top of the borrowing hierarchy in situations of language contact (Matras 1998). It has been further argued that an entire set of discourse markers in language A can

be gradually replaced by the corresponding lexical items from language B in the emerging, or already developed mixed codes (Matras 1998, 2000; Auer 1999), which brings us to the second part of the premise tested in this chapter.

The Russian lexical items analysed in this chapter appear throughout the corpus with virtually no competition from the Polish lexical elements and are exclusively employed in both Polish and Russian contexts, which further corroborates the premise tested in this chapter. This is particularly evident when it comes to the discourse particles ‘nu’ and ‘vot’. The fact that the two discourse particles are regularly employed by all the speakers in the data with numerous instances of use suggests that they cannot be regarded as sporadic occurrences, but rather permanent items stored in the lexicons of the individual speakers concerned. This finding also raises a further interesting question of whether this dynamic can be seen as a separation of Polish and Russian systems. The fact that some speakers are able to speak Russian without switching into Polish, as evident by a few examples from the present corpus, would suggest that they are able to keep at least one of the systems separate. However, as will be demonstrated in Chapter 6, there are numerous examples where Polish function words occur in utterances with Russian as the matrix language. Likewise, while some of the speakers would probably be able to inhibit the influence of Russian, especially on the lexical level, when talking to monolingual speakers of Standard Polish, judging from the density and complexity of language alternation attested in the present corpus, it would probably not be the case for the majority of speakers, which suggests that it would be plausible to regard the systems as at least partly unified.

This line of reasoning appears to be supported by psycholinguistic research. As the experimental evidence suggests, bilinguals are only partially able to inhibit the activation of one of their languages in selective word recognition tasks (Van Wijnendaele and Brysbaert 2002). Some researchers (e.g. de Bot 2000) therefore advocate the hypothesis that the systems are partly unified and partly separate, depending

on the degree of typological proximity between the languages and the speaker's proficiency levels in each. While links between elements of a single language tend to be stronger than between those in different languages, in the case of regular CS interlingual links may be as strong as intralingual ones. Bearing these observations in mind it is reasonable to assume that the Polish and Russian systems are likely to be unified to some extent since the two languages are typologically closely related and the speakers CS very regularly. Of course it should be noted that the reasoning provided here is based on speculation, rather than empirically grounded and would need to be further backed with experimental evidence.

While not as popular and frequent as the discourse particles 'nu' and 'vot', the three discourse markers analysed in the present chapter provide additional support for the Emerging Mixed Code/Fusion Hypothesis. This is particularly evident with lexical items 'koroče' and 'tipa', as both of them feature regularly in the discourse of the majority of speakers in the present corpus. Moreover, they can appear in both Polish and Russian contexts without any syntactic limitations or without triggering further switching, which suggests that, just like 'nu' and 'vot', they are no longer regarded as 'foreign' elements, but rather as an integral part of an emerging Polish-Russian mixed code. The same observations can be applied to the third discourse marker analysed in this chapter, namely the Russian 'prosto', though to a somewhat lesser degree. While its Polish cognate 'po prostu' does appear in the present corpus, it features predominately in the discourse of one speaker, which suggests that it is highly idiolectal. Therefore, this finding does not hinder the validity of the hypothesis, as for all the remaining speakers it is the Russian discourse marker that constitutes the default option. Bearing these observations in mind, one of crucial findings of the present chapter is that the Russian discourse markers 'koroče', 'tipa' and 'prosto' have been fused into an emerging Polish-Russian mixed code on a par with the discourse particles 'nu' and 'vot'.

Finally, the first premise of the Emerging Mixed Code/Fusion Hypothesis, tested in this chapter, can be further extended to include two Russian swearwords, ‘blin’ and ‘bljad’’, since they contribute to the organisation of the interaction and the structuring of verbal exchange, much like discourse markers. Just like discourse markers and discourse particles, these two swearwords appear with great frequency and regularity in the discourse of the majority of speakers without any immediate competing forms or any syntactic restrictions. Moreover, they can be employed as stand-alone items in the Polish base, which suggests that they are no longer regarded as ‘foreign’ elements, but rather as an integral part of an emerging Polish-Russian mixed code. Thus, it can be concluded that ‘blin’ and ‘bljad’’ have been ‘fused’ in an emerging Polish-Russian mixed code alongside Russian discourse markers and particles.

Chapter 5

Intraclausal Switches

As demonstrated in the existing research, language alternation of the insertional type is characterised by ad-hoc switches from the embedded language into the grammatical clause defined by the base/matrix language. These switches include single lexical items and larger constituents, which are made up of more than one lexical item but never surpass the clause (Auer 1999). Following Myers-Scotton (2002, 2005), these larger constituents will be referred to as ‘embedded language islands’ and together with single switches, they will be referred to as ‘intraclausal switches’. While the first premise of the Emerging Mixed Code/Fusion Hypothesis was tested in the previous chapter, the main aim of the present chapter is to test the remaining two premises. Recall from p.40 that the second premise states that intraclausal switching, consisting of single switches and embedded language islands is the preferred type of language alternation in the present data, while the third premise states that there is evidence of convergence of the regional variety of Polish towards Russian on the grammatical level, which manifests itself in the presence of structural and single-word calques.

In order to test the remaining two premises of the Emerging Mixed Code/Fusion Hypothesis the analysis presented in this chapter will focus on both lexical and grammatical manifestations of language mixing. Both of them arise as a result of the

same process, namely prolonged language contact, therefore it makes sense to analyse them in tandem. While the existing studies of CS often focus on either lexical or grammatical aspects of language alternation (rather than both), the approach taken in this study is to analyse these various contact phenomena together. This approach is particularly relevant when it comes to data that display CM since evidence of mixing on the structural level, which is often referred to as ‘convergence’, can further differentiate this type of language alternation from ‘classic’ CS. Myers-Scotton’s Matrix Language Frame Model, and the related 4-M Model will be employed as the theoretical framework for the data analysis.

As will be demonstrated in this chapter, the remaining two premises of the hypothesis are borne out by the Polish-Russian data. In line with the Emerging Mixed Code/Fusion hypothesis, it will be demonstrated that Russian intraclausal switches feature with great frequency and regularity in the discourse of all the speakers. Moreover, the imported lexical items get to keep their original Russian ending, which is the exact opposite of what Myers-Scotton’s model predicts. The implications of this finding will be further addressed in the concluding section of this chapter. Likewise, the data analysis demonstrates that the presence of structural and single-word calques is particularly prominent in the present corpus and shows that language alternation between Polish and Russian goes beyond the lexical level and can also be observed at the more abstract grammatical level. This finding is particularly important since it implies that we are dealing with a more advanced type of language alternation on the deeper structural level, as well as the lexical level.

The chapter is divided into four major parts. The first part gives an in-depth overview of the existing research on intraclausal switches, with a special focus on the distinction between single switches and lexical borrowings. It also introduces Myers-Scotton’s Matrix Language Frame Model, and the related 4-M Model, which constitute the theoretical framework for the present study. The second part focuses

on specific instances of single switches from Russian that are inserted into the Polish base, with a focus on nouns, verbs and adjectives, while the third part investigates larger constituents, known as ‘embedded language islands’. In addition to intraclausal switches, two types of structural borrowing, namely semantic borrowings and calques, are analysed in the fourth part of this chapter. The implications of the data analysis are discussed in the concluding section.

5.1 Switches or Borrowings?

The existing research suggests that one of the most common outcomes of the prolonged contact of several languages is the insertion of single lexical items from one language into the base of another language (Poplack et al. 1988; Matras 2009). These lexical items may feature in the discourse of only one speaker, or they may spread into the discourse of several speakers and over time become regularly used within the speech community. Moreover, they may appear unintegrated in the base language, or they may display various degrees of phonological and morphological alterations that have been made in accordance with the patterns characteristic of the base language. Since imported lexical items may differ significantly in terms of their distribution and degree of integration, a distinction between ‘single switches’ and ‘borrowings’ has been proposed in order to differentiate between spontaneous ad-hoc insertions on the one hand, and more established lexical items on the other hand (Poplack et al. 1988).

The distinction between single switches and lexical borrowings has received considerable attention in the existing literature. While different criteria have been put forward in order to differentiate between the two processes, consensus on this matter is yet to be reached. One of the most common criteria is the degree of phonological, morphological and syntactic integration into the recipient language (Poplack 1980;

Poplack and Meechan 1998). As Poplack and Meechan explain (1998: 127), CS involves alternation between two language systems, which means that single-word switches should display little or no integration into another language. On the other hand, the process of lexical borrowing refers to the incorporation of a lexical item from one language into another, with only the recipient system operative. What follows is that for a lexical item to be classified as a switch, it must retain the morphology, syntax and phonology of the donor language. However, if the lexical item assumes the morphological, syntactic, and often phonological identity of the recipient language, then it can be regarded as an established borrowing.

In addition to morpho-syntactic integration, other criteria that have been proposed include frequency of use, wide spread across the speech community, regularity of occurrence and lack of equivalent in the receiving language. Poplack (2004: 74) specifies that a lexical item that fulfils the criteria for being classified as an established borrowing will also be recurrent and widespread. However, it should be noted that the latter may prove problematic to verify. According to Auer (2007: 329), “frequency of usage is nothing but a very superficial generalisation over individual usage patterns.” Some lexical items may be highly idiolectal and appear rather frequently in the discourse of some speakers, whereas other speakers may not even be familiar with a given lexical form. Moreover, as Matras (2009: 106) adds, there is no uniform standard according to which a form’s frequency or the degree of spread across the community could be assessed. Instead, he proposes regularity of occurrence as an additional criteria, which is not measured in terms of frequency, but rather means that a lexical item “is independent of any contextual selection constraints and so is deemed appropriate in whichever language context it is being activated” (2009: 113), which suggests that it is also used as a “default expression, often a designation for a unique referent or a grammatical marker” (Matras 2009: 132).

Although the criterion based on the morpho-syntactic, and in some instances also

phonological integration has been advocated by Poplack and associates for a number of years (Poplack et al. 1988; Poplack and Meechan 1998), it has been proven to be insufficient when it comes to distinguishing between borrowings and switches. First of all, this criterion is not relevant in every context since it is very problematic to assess structural integration in languages that assign little inflectional morphology. Secondly, some lexical items remain only partially integrated in the recipient language, e.g. the English plural -a in ‘phenomena’. This suggests that structural integration itself is in fact subject to degree. Third, integration conventions may be highly idiolectal and therefore vary across individual speakers or from one bilingual speech community to another (Matras 2009: 113).

The existing studies have also demonstrated that ad-hoc integration of single switches is in fact possible and is regarded as the norm in most languages, which can be explained by referring to psycholinguistic aspects of language production. According to some researchers (e.g. Garrett 1980), function words and inflectional markers are part of the syntactic frame and are therefore retrieved in a different manner than content words. They remain in their original place in the sentence while content morphemes are interchanged. Since insertional CS normally involves content words, it is logical to assume that relevant inflections can be easily assigned to the selected lexical elements as they stay in their specified slots. The example below, taken from Angermayer’s (2005) study of English-Russian CS (quoted in Auer 2007) demonstrates that English single lexical items inserted into the Russian base tend to be assigned Russian inflectional endings and thus be subject to morpho-syntactic integration as soon as they are employed, e.g.:

čto vy думаete nepravil’no s etim estimate-om?
 what you think wrong with this estimate-INSTR?

‘what do you think is wrong with this estimate?’

According to Angermayer (ibid), the English noun *estimate* is an example of a single

switch, and not an established borrowing in Russian. It is integrated into the Russian sentence on an ad-hoc basis by adding the instrumental case ending, thus ensuring that even a single switch adheres to the rules of Russian morpho-syntax.

Further attesting to the fact that there is no reliable way of making a distinction between the processes of single switching and lexical borrowing is the view held by a number of researchers, who believe that the processes of CS and borrowing can be situated along a continuum (Auer 2007; Matras 2009; Gardner-Chloros 2009). According to Auer (2007: 327), ad-hoc insertions/nonce borrowings are situated at one end, while established loanwords at the other end. However, there are many intermediate points along the continuum, which are difficult to define. As Backus explains (1992: 15), it is impossible to tell at which stage of the borrowing process a given lexical item is currently situated, since acceptability and integration criteria may vary from speaker to speaker, and from one speech community to another. While some speakers may see a given lexical item as an established borrowing, others will continue to treat it as a word belonging to a different language.

Finally, it should be borne in mind that the term ‘borrowing’ is an umbrella term which is employed to refer to several processes. According to Matras and Sakel (2007), borrowing can be further divided into ‘material borrowing’ and ‘structural borrowing’. Material borrowing refers to sound-meaning pairs, which most often include lexemes and are also known as ‘loanwords’, while structural borrowing refers to the copying of syntactic, morphological or semantic patterns, and includes calques and semantic borrowings (Haspelmath 2009: 39).

While the distinction between single switches and material borrowing is often addressed in classical studies of CS, other contact phenomena, such as calques and grammatical interference are ignored, even though all of these phenomena co-occur in bilingual corpora. Although lexical insertions operate on the surface level while calques belong to the abstract level, essentially both of them can be regarded as the outcome of

the same process, namely contact-induced influence of the embedded language on the matrix language. While the presence of grammatical interference is sometimes acknowledged in studies of CS (e.g. Doğruöz and Backus 2009), a coherent integrative framework that would study both phenomena together is yet to be developed. Bearing these observations in mind, it makes sense to enlarge the scope of analysis which is usually adopted in classical studies of CS and look at both lexical and grammatical outcomes in tandem.

As far as lexical borrowing is concerned, the criterion based on the degree of morpho-syntactic integration is likely to prove insufficient and difficult to implement in the Polish-Russian data, which raises a further important issue, namely the typological closeness between the two languages studied in this thesis. Polish and Russian are typologically closely related, thus they share not only lexical similarities in the form of homophones and cognates, but are also morphologically and structurally congruent to a very high degree. One of the consequences of the typological proximity is the degree to which borrowed, as well as switched elements can be marked with the matrix language morphology. As will be demonstrated in the remainder of this chapter, the local variety of Polish, which normally acts as the matrix language, easily adopts Russian lexical elements and freely adds relevant suffixes due to the high compatibility of the two morphological systems. While on the one hand, typological similarity makes it more difficult to apply the criterion based on morpho-syntactic integration to classify a given lexical item as a case of lexical borrowing, on the other hand, it facilitates switching, thus the second consequence is that structural similarities between Polish and Russian create favourable conditions not only for CS, but also complex mixing. The consequences of typological closeness will be further addressed in the concluding chapter of this thesis.

In order to classify a given lexical item as an instance of lexical borrowing it may be worth turning to the more traditional criteria, such as frequency of use and Matras'

criterion of regularity of occurrence, measured in terms of contextual independence and lack of competing equivalents. However, these criteria are quite difficult to implement since the present corpus is relatively small. These observations further corroborate the view expressed in some of the existing literature that the distinction between single switches and borrowings will never be reliable, hence the following section does not aim to make any categorical claims, but to situate a given lexical item as being closer to the switching or the borrowing end of the continuum, and to further prove the extent to which switching is prominent in the present data.

5.1.1 Lexical Borrowings/Loanwords in the Present Corpus

Following the survey of the existing literature presented above, it can be concluded that the distinction between single switches and instances of material borrowings/loanwords is a problematic issue. However, it is possible to make some tentative conclusions by analysing such factors as regularity of occurrence, measured in terms of contextual independence, and frequency of use. Bearing these criteria in mind it should be possible to situate a given lexical item closer to the borrowing- or switching-end of the continuum.

As far as the present data are concerned, it should be stated at the very beginning that the process of material borrowing is nowhere near as productive as that of switching, for there are only two lexical items that can be singled out as potential candidates to be classified as lexical loanwords. They are ‘sumka’ - *bag* and ‘maszyna’ - *car*.

The main difficulty with establishing the status of the noun ‘sumka’ is related to the fact that it does not feature very frequently in the corpus. It is employed by three different speakers on six occasions. Its Standard Polish equivalent ‘torba’ does not appear in the present corpus, which to some extent suggests that this lexical item is selected as the default option independently of the context in which it appears, and thus fulfils one of the criteria to be classified as a borrowing. An example of its

contextual usage can be found in Extract 32:

Extract 32

- (1) A: sumka bierz ze sobo
bag take with self
'take the bag with you'

In the above extract the noun 'sumka' is employed in line with the rules of regional Polish morpho-syntax. It is assigned the accusative ending '-a', characteristic of the feminine nouns in the regional variety of Polish, though as will be demonstrated later in this chapter, single switches can also be integrated in a similar way. While the lack of immediate competing equivalents from Polish suggests that it could indeed be an instance of material borrowing, its phonological integration is more problematic to assess. The reason for this is the fact that there are no instances of 'sumka' used in the entirely Russian context, which would allow to compare the realisation of the stressed vowel /u/ and the unstressed vowel /a/ in the Russian context and in the Polish context, and thus assess its degree of integration in the Polish base.

An even better candidate is the noun 'maszyna' - *car*, which is employed on 32 occasions by ten different speakers. An example of its contextual usage can be found in Extract 33:

Extract 33

- (1) A: dla swojej dziewczyny taka maszyna kupiłby?
for your girlfriend such car would buy
'would you buy such a car for your girlfriend?'

As with the previous example, the noun 'maszyna' is assigned the accusative ending '-a', which is characteristic of the regional variety of Polish spoken in Lithuania. However, this cannot be regarded as a convincing piece of evidence since single switches from Russian can also be integrated on an ad-hoc basis in exactly the same

way. Unlike ‘sumka’, the noun ‘maszyna’ appears in the corpus together with its Standard Polish equivalent ‘samochód’, which is employed on a total of 22 occasions by nine speakers. The distribution of both lexical items in the Polish base is summarised in Table 5.1 below.

Speaker	‘maszyna’	‘samochód’
A	-	8
B	-	1
C	2	-
D	1	3
F	1	-
G	-	1
N	4	1
M	1	3
Q	1	3
R	-	1
S	5	-
T	10	1
U	3	-
V	4	-

Table 5.1: Distribution of ‘maszyna/samochód’ in the Polish base

According to the table, overall ‘maszyna’ is employed more often than ‘samochód’, but there are five speakers who employ both lexical items in their discourse simultaneously. When a loanword and a non-borrowed native synonym continue to co-exist, which seems to be the case with ‘maszyna’ and ‘samochód’ in the present data, they may become subject to stylistic specialisation, which affects both the borrowed lexical item and its non-borrowed counterpart. In German, for example, ‘Fernsprecher’ and ‘Telefon’, are both employed with the meaning of *telephone*, but the former is normally

reserved for official or institutional contexts. With regard to the Polish-Russian data, at this stage it seems plausible to treat ‘maszyna’ as situated closer to the borrowing end of the continuum, though it is too early to say whether there is any stylistic specialisation between the native ‘samochód’ and the borrowed ‘maszyna’, which is largely due to the fact that the present corpus is too small to draw any substantial conclusions.

With only two lexical items identified as potential loanwords, it appears that the process of material borrowing is not very productive in the present data. As will be demonstrated in the following sections, the majority of Russian lexical items are single switches with very limited or no recurrence, though this of course could be related to the fact that the present corpus is too limited to make a proper assessment. Preference for switching over material borrowing could be conditioned by the fact that frequent mixing of Polish and Russian is accepted within the Polish speech community, which permits speakers to insert ad-hoc lexical items from Russian without the need to borrow them long-term. The fact that Polish and Russian are two closely related languages facilitates this process even further, since in the majority of instances morpho-syntactic integration of Russian lexical items is not even required. In the examples of ‘maszyna’ and ‘sumka’ that feature in the present corpus, the Russian morphemes either overlap entirely with the Polish ones, or they are assigned the inflectional ending ‘-a’ in the accusative case, which is characteristic of the regional variety of Polish. However, as will be demonstrated in section 5.3.1 this ending also tends to be assigned to those lexical items that can be classified as single switches.

Unlike material borrowing, intraclausal switching of the insertional type, which involves single switches and embedded language islands is extremely productive in the present data. One of the most comprehensive theoretical frameworks that have been developed to analyse language alternation of this kind is Myers-Scotton’s Matrix

Language Frame model (1993, 2002). It was later supplemented with the additional 4-M model, which focuses specifically on the four types of morphemes and their distribution in CS. Both of these models will constitute the analytical framework for the present chapter and are presented in greater detail in the following section.

5.2 Matrix Language Frame Model and the 4-M Model

When it comes to analysing the grammatical outcomes of language alternation, Myers-Scotton's Matrix Language Frame model (1993, 2002) (henceforth MLF) is often cited as the most comprehensive and influential framework in this field (e.g. Clyne 2003; Deuchar 2005). The main aim of the MLF model is to provide a set of criteria for identifying cases of 'classic' CS, which can be defined as a type of language alternation that "includes elements from two (or more) language varieties in the same clause, but only one of these varieties is the source of the morphosyntactic frame for the clause" (Myers-Scotton 2002: 241). The term 'morphosyntactic frame' refers to all the abstract grammatical requirements, which ensure that the frame is well-formed, namely word order, morpheme order and the inflectional morphemes.

As the definition of 'classic' CS explains, the MLF model only deals with intraclausal switching, i.e. language alternation which takes place within the clause, which according to Myers-Scotton (2005: 241), is the best unit of analysis for the bilingual data. While switching can also be interclausal, i.e. have one clause from one language, and the other clause from another language, it does not show the two languages actually in contact, since the grammars of the two languages do not interact the same way as they do when intraclausal switching is involved (ibid). Moreover, the status of the clause is clearer than that of the sentence, which can be particularly problematic to define when dealing with spoken language data. The clause, on the other hand, can

be defined unambiguously in terms of phrase structure as a complementiser or an element in the specifier position followed by an IP (2002: 55).

It should be borne in mind that the MLF model is concerned only with grammatical aspects of CS and therefore it does not attempt to analyse any discourse-related functions of CS, which are often performed when switching takes places between clauses or between turns. In other words, the MLF model only investigates CS of the insertional type and does not deal with CS of the alternational type. While Myers-Scotton's model will be employed as a theoretical framework for analysing instances of the former, the latter type of switching will also be addressed in Chapter 6. Consequently, as stated in the literature review, the notion of 'classic' CS adopted in this thesis does also, unlike Myers-Scotton, accept interclausal switches to be relevant phenomena of CS.

Following Joshi (1985), the first premise of the MLF model is that the frame-building language, i.e. the Matrix Language, and the language that the speaker switches to, i.e. the Embedded Language, do not participate equally in the constituent structure. The second premise is that not all morpheme types can come equally from the Matrix Language and the Embedded Language. The third premise is that both languages are activated when a speaker engages in CS, though it is the Matrix Language that is activated more. On the basis of these premises, Myers-Scotton (2002: 244) formulates two principles which apply to mixed constituents, i.e. constituents that include morphemes from both languages. The first one is the Morpheme Order Principle, which is formulated as follows:

In mixed constituents consisting of at least one Embedded Language word and any number of Matrix Language morphemes, surface word (and morpheme) order will be that of the Matrix Language.

The second principle, called the System Morpheme Principle, states the following:

In Matrix Language+Embedded Language Constituents, all system morphemes which have grammatical relations external to their head constituents will come from the Matrix Language, i.e. the so-called ‘late’ system morphemes under the 4-M Model.

The central idea of the MLF model, and the related 4-M Model is the division of morphemes into content and system morphemes. Content morphemes can be defined as those that assign and receive thematic roles and include nominal and verbal stems. In contrast with content morphemes, system morphemes do not receive or assign thematic roles. Prototypical system morphemes include all affixes (bound morphemes) and some function words that stand alone (e.g. determiners and clitics). Since system morphemes are more varied than content morphemes, Myers-Scotton and Jake (2000, 2001) propose an additional model, called the ‘4-M Model’, where they further divide system morphemes into several types based on how and when they are activated in the process of language production. Thus, content morphemes and early system morphemes are ‘conceptually activated’ by the speaker’s intentions. This activation happens at the level of the mental lexicon, i.e. the first level, hence the label ‘early’. Early system morphemes include such meaningful elements as plural markings, determiners, and prepositions that change the meaning of phrasal verbs (Myers-Scotton 2002: 268).

The other two types of system morphemes are called ‘late’ because according to the 4-M Model, they are not activated until a later production level when larger constituents are assembled. The late system morphemes are further divided into two types, namely the ‘bridge system morphemes’ and the ‘outsider system morphemes’. As the name implies, bridge system morphemes occur between phrases that make up a larger constituent. Some examples of bridge morphemes include the possessive *'s* in English or the preposition *of*, which signal the relationship between the possessor noun and the element that is possessed. The other type of late system morpheme is

the outsider system morpheme. The form of this morpheme depends on information outside the element with which it occurs and therefore outside its immediate phrase. That information comes from an element in another constituent or from the discourse as a whole. An example of an outsider morpheme is the element that shows subject-verb agreement, since the form of the agreement marker depends on the subject, e.g. the suffix *-s* in English. Another example of outsider morphemes are case markers, for they are often determined by verbs and prepositions, and therefore their form depends on information outside the phrase in which they occur (Myers-Scotton 2002: 269–270).

The MLF model, as well as the related 4-M model appear to be particularly appealing for analysing instances of intraclausal switches between such highly inflected languages as Polish and Russian. The division between early and late system morphemes, proposed under the 4-M model, permits to provide more accurate explanations for what happens in intraclausal switching where case markers are involved. This view has been endorsed by some of the existing literature, most notably Halmari who studied English-Finnish CS and concluded that “[the MLF model] seems to suit well to explain data from morphologically complex languages” (1997: 68). Moreover, the MLF model also provides additional basis for making a distinction between ‘classic’ CS on the one hand and ‘composite’ CS (referred to as CM in the present study) on the other hand, which falls in line with one of the hypotheses tested in this thesis. As stated on p.36 in the literature review section, ‘composite’ switching can be regarded as a combination of CS and convergence. Recall from p.36-37 that in ‘composite’ CS one of the participating languages contributes most of the morpho-syntactic structure. However, the other participating language also contributes some of the abstract structure underlying surface forms in the clause. What this means for language mixing is that in ‘composite’ switching not only surface-level morphemes originate from both languages, but also the abstract structure of the frame itself, hence the presence of convergence can be regarded as a significant piece of evidence

for the Emerging Mixed Code Hypothesis.

However, as will be further demonstrated in the following section, the MLF model does have some major drawbacks. One of Myers-Scotton's claims is that intrasentential CS is governed by abstract principles which apply universally (1995: 234). Her model is therefore valid across different speech communities since it is based on 'pre-syntactic' principles (1995: 251). In other words, the MLF is a production-based model, which is concerned with the abstract structure of the language, thus the underlying processes that it is based on do not depend on the participating languages. The apparently universal appeal of the MLF model has been received with scepticism by some of the researchers. According to Jacobson (1998: 61), it is unlikely that the model lends itself well to postulating global constraints that exclude all violations.

Likewise, some researchers found it problematic to apply the MLF model to CS data which involve typologically closely related languages. Berruto (2005: 89), who analyses language alternation between Italian and Italo-Romance dialects, found that in his corpus the System Morpheme Principle is often contradicted, which leads him to the conclusion that "the general principle that the Matrix Language alone supplies the system morphemes within mixed constituents cannot be sustained" (2005: 91). He further questions the validity of the matrix language concept itself, for "it cannot be convincingly applied to cases where the speakers continually switch to and fro between two linguistic systems" (2005: 92). Similar issues are reported by Franceschini (1998: 58-59) who shows that in CS between Standard Italian and Lombard dialect embedded language system morphemes are permitted.

It should be noted that the two studies quoted above do not make any reference to the 4-M Model, which was developed much later than the MLF model itself and which introduces a further division of system morphemes into early and late. Following this modification, Russian lexical items embedded into the Polish base must be assigned Polish late system morphemes, such as case markers in nouns, though they need not

be assigned bridge system morphemes. The opposite case scenario, where Russian lexical items retain their late system morphemes when embedded into the Polish base would be considered as a significant deviation.

However, as will be demonstrated in this chapter, even with the additional subdivision into the early and late system morphemes, the MLF model is problematic to apply to such typologically closely related languages as Polish and Russian. The two languages already share some inflectional endings, which means that there are instances when no overt morpho-syntactic alterations are required to ensure that the lexical item is an integral part of the Polish base. Moreover, a number of Polish and Russian morphemes are very similar. For example, the masculine and neuter instrumental case ending in Russian is ‘-om’ (/om/), while its Polish cognate is ‘-em’ (/em/). While these two forms are distinct morphemes, they nevertheless resemble each other to a great extent, which means that the use of the Russian or the Polish morphological form does not have any effect on the grammatical make-up of the sentence.

Bearing the above observations in mind, it comes as no surprise that in the majority of instances in the present data the so-called ‘outsider late systems morphemes’ in switched lexical items remain Russian. Contrary to Myers-Scotton’s model, it is actually surprising when the imported lexical item is assigned a late system morpheme of the receiving language. As far as nouns are concerned, the majority of Russian lexical items appear in their original form, the only exception being feminine nouns which are assigned the ending ‘-a’ in the accusative case, characteristic of the regional variety of Polish spoken in Lithuania. Likewise, Russian verbs tend to appear in their original form and thus keep Russian late system morphemes, or they undergo a complete restructuring and surface as seemingly Polish forms, which will be discussed as the process of calquing in section 5.5.2. Similarly, the majority of adjectives keep their original bridge and late system morphemes, though in some feminine adjectives they tend to be replaced with shorter endings in analogy with Polish forms. Each

of these major word classes will be addressed separately in the next section and the implications of the analysis presented below will be addressed in the concluding section of this chapter.

5.3 Single Switches in the Present Corpus

As stated in the introduction, the present corpus consists of 109,722 orthographic words, out of which there are 89,015 Polish words, 20,093 Russian words and 614 Lithuanian words. Intraclausal switches, which consist of single switches and embedded language islands amount to 3,521 tokens (39% of the total number of all types of switches). This suggests that after discourse markers, they are the second preferred type of language alternation. In contrast, interclausal switches, which are addressed in greater detail in the next chapter, are employed only on 322 occasions (3% of all switches). Such great quantitative discrepancy demonstrates that intraclausal switches constitute the preferred type of language alternation in the present data and are outnumbered only by discourse markers. However, a point needs to be made with regard to these two types of language alternation.

Discourse markers are made up of a set inventory of recurrent tokens, i.e. the same limited set of lexical items is inserted into the Polish base over and over again. As will be further explained in the remainder of this chapter, intraclausal switches are non-recurrent ad-hoc insertions, which in the majority of instances are employed on one occasion only. This suggests that the two types of language alternation are qualitatively different, since discourse markers can essentially be treated as a type of material borrowing. Therefore, in order to establish the preferred type of switching in the present data it is more reasonable to compare the distribution of intraclausal switches against the distribution of interclausal switches, and treat discourse markers as a separate category, situated between the switching and borrowing end of the

continuum. Following this line of reasoning, intraclausal switching would be classified as the preferred type of ‘proper’ language alternation attested in the present data.

This section is concerned specifically with single switches, which amount to a total of 2,605 tokens. The analysis will focus on three major word classes, which are employed most frequently as single switches in a Polish base. These are made up of 1,484 tokens of nouns, 680 tokens of verbs and 314 tokens of adjectives. Ideally, in order to make an appropriate comparison, these figures would need to be contrasted with the number of tokens for the corresponding classes of Polish lexical items that appear in the data, e.g. the number of Russian nouns that appear as single switches against the total number of Polish nouns. However, at this stage it is not possible to make such a comparison due to one of the drawbacks of the present corpus, namely the fact that, with the exception of switches, it is untagged and therefore part-of-speech tags for each lexical item have not been specified. As a result, some types of frequency information, such as the number of tokens for each lexical category, cannot be extracted at the moment and therefore it is not possible to make any further generalisations about speakers’ preferences with regard to switching at given lexical categories. While it appears that speakers prefer to switch at nouns, this could only be backed up empirically if the number of switched Russian nouns was contrasted against the number of Polish nouns. Corpus annotation therefore remains one of the issues that could be addressed during future research.

The initial observation that emerges with regard to Russian single switches is that most of them can be described as non-recurrent. In the majority of instances a given lexical item is used only once and does not reappear in the remaining discourse of the same or other speakers. Of course this could be due to the fact that the present corpus is too small and therefore too limited to make a proper assessment of the frequency and distribution of single lexical items. Nonetheless, the very fact that Russian lexical items are often used only once may suggest that we are dealing with

a prototypical case of strong bilingualism and heavy CM.

In addition to being non-recurrent, the majority of single switches appear to be idiosyncratic and stylistically neutral, i.e. they come from random semantic categories and it would be difficult to group them together under one label, e.g.: ‘volny’ – *waves*, ‘učeba’ – *studies*, ‘krajnost’ – *extremity*, ‘mečta’ – *dream*, ‘nažat’ – *to press*, ‘otpraviti’ – *to send*, ‘ponjatnyj’ – ‘clear’, ‘semejnyj’ – *family-oriented*. Nevertheless, it is possible to make several initial generalisations that emerge in terms of the semantic and grammatical profile of some of these lexical items. First of all, there is a sizeable group of highly specialised technical terms, which feature rather prominently in conversations between male respondents, which can also be linked to the topic of their conversations. Quite often these are lexical items referring to parts of a car, e.g.: ‘farkop’ – *hitch*, ‘peredača’ – *gear*, ‘sceplenie’ – *clutch*, ‘tormoza’ – *brake*, or to construction, e.g.: ‘lebedka’ – *hoist*, ‘pogruzčik, bobkat’ – *loader*, ‘osnova’ – *foundation*, ‘betonosmesitel’ – *concrete mixer*. Secondly, in terms of their stylistic nature a lot of the switched lexical items can be described as informal. These lexical items tend to be highly expressive and therefore they help speakers modify the propositional content of the utterance. Some examples include such nouns as ‘musora’ – *cops*, ‘bugaj’ – lit. ‘bull’ – *hench*, ‘čuvak’ – *dude*, verbs such as ‘tusovat’sja’ – *to party*, ‘shljat’sja’ – *hang around*, ‘muryžit’ – *to put off*, and adjectives such as ‘ofigennyj’ – *awesome*, ‘čoknutyj’ – *crazy*.

As far as the grammatical profile of switched lexical items is concerned, one generalisation can be made with regard to the category of nouns. As stated above, there is a sizeable proportion of nouns that are highly informal and colloquial in nature. This can be partially achieved with the addition of diminutive, augmentative or simply expressive suffixes, which are employed to add different hues of meaning to the original lexical items. According to Offord and Gogolitsyna (2005: 277), in the Russian language various suffixes may be added to express the speaker’s attitude, which can

range from affectionate, tender, affective to scornful, ironic, and disparaging. In the majority of cases one and the same suffix performs a two-fold function, i.e. indicates the size of an object and expresses the speaker's attitude towards it. Some examples of nouns with augmentative and/or expressive suffixes include 'zakuson' – *snack*, 'muzon' – *music*, 'domina' – *house*, 'audjuha' – *Audi*, while examples of nouns with diminutive suffixes include 'dedulek' – *grandpa*, 'dirik' – *director*, 'batjok' – *dad*, 'babul'ka' – *granny*. In all of these instances the speakers' attitude can be described as playful, ironic or disparaging.

As stated in the introduction to this section, the findings presented above are based on observation, rather than systematic and detailed analysis of the corpus. The extent to which they are valid for the present data needs further investigation, which cannot be carried out in the scope of this chapter. Since there is a great inventory of switched tokens involved, it would be extremely time-consuming to perform a quantitative analysis for each of them, hence the information presented above is intended to give the reader a preliminary overview of the corpus and the nature of single switches which will be discussed in greater detail in the remainder of this chapter.

5.3.1 Nouns

With a total of 1,484 tokens, nouns constitute by far the largest group of single switches in the present corpus. Recall from p.118 that according to Myers-Scotton's System Morpheme Principle, in mixed constituents all system morphemes which have grammatical relations external to their head constituents will come from the matrix language, i.e. the so-called 'late' system morphemes under the 4-M Model. As far as nouns are concerned, the relevant 'late' system morphemes are case markers, for they are assigned by verbs and prepositions, and therefore their form depends on information outside the phrase in which they occur (Myers-Scotton 2002: 269–270). Therefore, following the MLF model, it can be expected that Polish 'late' system

morphemes, or case markers as far as nouns are concerned, will be assigned to Russian switches. However, the data show that three possibilities can in fact be singled out. The first possibility are Russian nouns that have exactly the same case ending in Polish, the second possibility are Russian nouns which keep their original Russian inflection, and the third possibility are Russian nouns integrated into the Polish base by means of a Polish case ending. In the first and third instances, Russian nouns appear to be morpho-syntactically integrated into the Polish base/matrix language, while in the second instance their presence does not render the base/matrix language ungrammatical, although there is no overt manifestation of their morpho-syntactic integration. A detailed quantitative breakdown of tokens for each case and gender combination can be found in Table 5.2 below.

The first possibility are nouns with zero ending (\emptyset). Some examples can be found in Extracts 34 and 35:

Extract 34

- (1) F: skrócili jego _[RUS] **rejs** (0.2) i drugiego człowieka posłali
 cut his journey (0.2) and another person sent
 ‘[they] cut his journey (0.2) and sent another person instead’

In the above example the Russian noun ‘rejs’ appears in the accusative singular case. In both Polish and Russian the accusative case of the inanimate masculine nouns takes zero ending, thus it is very easy to insert the Russian noun into the base/matrix language since no morphological or syntactic adjustments need to be made for a noun to be integrated.

This Russian noun does have an immediate Polish cognate ‘rejs’. While the morpho-syntactic characteristics of the noun make it difficult to tell whether it is the Russian or the Polish variant that is being used, its phonological shape indicates that it is clearly a Russian noun. This becomes particularly evident when the realisation of the

Zero ending (\emptyset)	Ambiguous ending	Unambiguously Russian ending	Unambiguously Polish ending
Total: 376 tokens	Total: 949 tokens	Total: 66 tokens	Total: 93 tokens
<ul style="list-style-type: none"> • NOM sg masc – 129 • ACC sg masc – 191 • NOM sg fem – 14 • GEN pl fem – 42 	<ul style="list-style-type: none"> • GEN sg masc – 370 • DAT sg masc – 74 • NOM sg fem – 27 • GEN sg fem – 158 • DAT sg fem – 5 • LOC sg fem – 10 • GEN sg neut – 16 • DAT sg neut – 5 • NOM pl masc – 79 • INST pl masc – 18 • LOC pl masc – 28 • NOM pl fem – 25 • ACC pl fem – 91 • NOM pl neut – 28 • INST pl neut – 15 	<ul style="list-style-type: none"> • INST sg masc – 37 • INST sg fem – 24 • ACC sg masc – 5 	<ul style="list-style-type: none"> • ACC sg fem – 87 • GEN pl fem – 6

Table 5.2: Breakdown of single switched nouns for each case/gender

consonantal phoneme /r/ is compared. In Russian consonants are palatalised before the front vowel /e/, hence /r/ is realised as [rʲ]. When comparing the realisation of /r/ in speaker F's pronunciation, palatalisation is very prominent and leaves little doubt as to the fact that it is the Russian lexical item that is used. Moreover, the quality of the vowel /e/ is very different in Polish and in Russian. The vowel in the Polish 'rejs' could be transcribed as [ɛ], whereas the vowel in the Russian 'rejs' could be represented as [e]. The vowel in speaker F's pronunciation sounds definitely closer to the latter than the former. This means that the Russian noun is integrated into the base/matrix language, but it has not been phonetically assimilated into the host language.

Extract 35 contains another example:

Extract 35

- (1) M: _[RUS] **kapot** trzeba otkrywać
 bonnet must open
 '[One] must open the bonnet'

In the example above the Russian noun 'kapot' is employed in the accusative singular case. It has no overt case inflection and appears in the same form as it would in the nominative case. Since it is integrated into the base/matrix language as far as morphology and syntax are concerned, its status of a single switch can only be confirmed by analysing its phonological features. The lexical item 'kapot' has no immediate cognate in Polish, which means that it is definitely a Russian noun. If it were to be integrated into the Polish base, one of the immediate prosodic adjustments would be a change of stress from the last syllable to the initial syllable since Polish has fixed stress on the penultimate syllable. In the example above the original Russian stress is retained, which is a strong indication that the lexical item behaves in line with the rules of Russian phonology.

While Extracts 34 and 35 exemplify Russian nouns with a zero ending, the following

extract contains an example of a noun whose ending overlaps in Polish and Russian. When this happens, the noun appears to be integrated into the base/matrix language. However, the morphological overlap makes it impossible to tell which set of morpho-syntactic rules is being followed. An example can be found in Extract 36:

Extract 36

- (1) M: _[RUS] **soljarki** tyle zsmarnował
 diesel so much wasted
 ‘[I] wasted so much diesel’

In the example above the Russian noun ‘soljarki’ is feminine and is employed in the genitive singular case. The suffix ‘-i’, which is assigned to mark this case in feminine singular nouns coincides in both Polish and Russian, e.g.: ‘soljarka’ [NOM sg] – ‘soljarki’ [GEN sg], ‘drukarka’ [NOM sg] – ‘drukarki’ [GEN sg]. As with the previous two examples, the morphological form alone is not sufficient to assess the degree of integration into the matrix language, thus it is important to take the phonological characteristics into account.

When used in a Polish base the noun ‘soljarki’ seems to be pronounced in a similar manner to how it is normally pronounced in Russian. This is particularly evident in speaker M’s pronunciation of the ‘lja’ sequence, where the Russian phoneme /l/ is clearly palatalised and can be represented as [lʲa]. The phoneme /l/ is never palatalised before back vowels in Standard Polish, thus if it were to follow the rules of Polish phonology the sequence would be realised as [la]. Moreover, speaker M pronounces the unstressed phoneme /o/ in the pretonic syllable as it would normally be in Russian, i.e. as [ʌ]. This suggests that the Russian lexical item is imported into the Polish base with its distinct Russian pronunciation and morphology.

The second option is when a Russian noun is employed with an unambiguously Russian inflectional ending. When this happens, the noun displays no overt integration

into the base/matrix language since it retains all the morphemes of the embedded language, though there may still be phonological integration. In the overwhelming majority of cases, namely 61 tokens throughout the corpus, these inflections are those of the instrumental case (feminine and masculine), with occasional instances of the locative case endings, which will be discussed in the analysis of embedded language islands in section 5.4. An example of the Russian instrumental case morpheme employed in the Polish base can be found in Extract 37:

Extract 37

- (1) G: _[RUS] w kinach jest (.) _[RUS] **mul'tiki** (.) po litewsku
in cinemas is (.) cartoons (.) in Lithuanian
‘in the cinemas there are (.) cartoons (.) in Lithuanian’
- (2) F: _[RUS] **da?**
yes?
‘really?’
- (3) G: z litewskim _[RUS] **perevodom**
with Lithuanian translation-INST
‘with Lithuanian translation’

The lexical item ‘perevodom’ is imported into the Polish base along with its Russian inflectional morpheme and its original Russian pronunciation. While the articulation of /o/ may not be the most reliable feature for distinguishing between the Polish and the Russian pronunciation, there are other phonological criteria that can be employed. The fact that ‘perevodom’ retains its Russian articulation becomes particularly evident when the realisation of the consonantal phonemes /r/ and /p/ is compared. In Russian consonants are palatalised before the front vowel /e/, hence /p/ is realised as [pʲ] and /r/ is realised as [rʲ].

The third option is when Russian nouns are integrated into the Polish base by means of adding a Polish suffix. The most common group of nouns that are subject to this

type of integration are Russian feminine nouns that are employed in the accusative singular case in the Polish base, with a total number of 87 tokens across the corpus. The addition of the accusative singular Polish inflectional ending in feminine nouns appears with great regularity and can be regarded as a rule, since it can be observed across all speakers and in all instances of singly-occurring Russian nouns. It is difficult to tell what contributed to this relative success of the Polish feminine accusative ending other than the fact that it has become one of the staple features of the variety of Polish spoken in Lithuania and has been reported in the existing literature on numerous occasions (cf. Kurzowa 1993; Karaś 2009).

The second group of nouns that are sometimes subject to morphological integration are Russian feminine nouns used in the genitive plural case, though much less frequently (a total of six tokens). In Standard feminine and some neuter nouns, employed in the genitive plural are normally assigned zero ending (\emptyset), e.g.: [PL] ‘gra’ [NOM sg] – gier [GEN pl] – (*of*) *games*, [PL] ‘miasto’ [NOM sg] – ‘miast’ [GEN pl] – (*of*) *towns*, [RUS] ‘kassa’ [NOM sg] – ‘kass’ - (*of*) *tills*, [RUS] ‘vino’ [NOM sg] – ‘vin’ [GEN pl] – (*of*) *wines*. What happens in the regional variety of Polish spoken in Lithuania is that all nouns tend to be assigned the genitive masculine plural ending ‘-ów’. Some examples from the corpus include such lexical items as ‘miejsców’, instead of Standard Polish ‘miejsc’ – (*of*) *places*, ‘dziewczynów’, instead of Standard Polish ‘dziewczyn’ – (*of*) *girls*, and ‘grów’, instead of Standard Polish ‘gier’ – (*of*) *games*. Both case markers, the accusative feminine singular ‘-a’ and the genitive masculine plural ending ‘-ów’, can be classified as outsider late system morphemes since they are assigned either by the verb, or the adnominal genitive. Extracts 38–41 provide several examples of when these case markers are assigned to Russian single switches:

Extract 38

- (1) U: trąba zdjął (.)_[RUS] **zašita** zdjął
 horn removed (.) protection removed

‘[he] removed the horn (.) removed the protection’

Extract 38 exemplifies two feminine singular nouns employed in the accusative case, the Polish ‘trąba’ and the Russian ‘zašita’, whose case is assigned by the verb ‘zdjął’. Unlike in Standard Polish, where the corresponding suffix is ‘-ę’ (‘trąb-ę’), in the regional variety of Polish the accusative singular feminine form is the same as the nominative. This feature has been identified as one of the grammatical characteristics of the Polish spoken in Lithuania (Karaś 2009), where the Standard Polish ending ‘-ę’ is realised as ‘-a’. Therefore, it appears that Russian feminine singular noun ‘zašita’ is employed in analogy with the Polish one, and is integrated into the Polish base by adding the suffix ‘-a’.

While it may be argued that in Extract 38 the addition of the inflectional ending ‘-a’ to the Russian noun is influenced by the immediate presence of the Polish noun which is used in exactly the same form, Extract 39 demonstrates that this rule also applies to Russian nouns that appear without any influence from Polish lexical items:

Extract 39

- (1) Q: będzim _[RUS] **kryša** mieniać
will be roof replace
‘[we] will be replacing the roof’

Despite their apparent morphological integration, Russian lexical items in Extracts 38 and 39 nevertheless behave as prototypical single switches. This is particularly evident in Extract 38 where the lexical item remains phonologically unintegrated and retains its Russian pronunciation. The sequence ‘ši’ in ‘zašita’ features a long voiceless palato-alveolar fricative, which can be represented as [çç], which is characteristic of Russian, rather than Polish. It should be noted that Polish does have a similar sound sequence [ʃtʃ] (spelled ‘szcz’), which can be found in Russian-Polish cognates, e.g.: ‘eše’ – ‘jeszcze’ – *more*, ‘borš’ – ‘barszcz’ – *borscht*. However, qualitatively it is quite

different from the Russian [ɕɕ], most notably in its length of articulation and degree of palatalisation. If ‘zašita’ were adapted into Polish on the phonological level, most likely it would take the surface form [zaʃtʃiʦita], which is definitely not the case as far as Extract 38 is concerned.

As far as the lexical item ‘kryša’ in Extract 39 is concerned, the most prominent element for evaluating the degree of phonological integration is the phoneme /i/, which is articulated closer to the roof of the mouth in Russian than in Polish. While this difference in articulation is relatively subtle, the vowel that appears in speaker Q’s articulation of ‘krysha’ resembles the Russian variant, rather than the Polish one.

While much less frequent than the accusative ending ‘-a’, some regularity can be observed in plural genitive feminine and neuter forms, which are sometimes assigned the regional Polish inflectional ending ‘-ów’. Some examples can be found in Extracts 40 and 41 below:

Extract 40

- (1) U: tylko nie [kupu] [RUS] **bosonożki** (.) ni lubia [RUS] **bosonożk-ów**
 only not [buy] sandals (.) not like sandals
 ‘only don’t [buy] sandals (.) I don’t like sandals’

In the first instance ‘bosonożki’ is used in the accusative plural case with the ending /ɪ/, which corresponds in Polish and Russian, thus it is difficult to tell which morphological system is being followed here. In the second clause the speaker clearly uses the genitive case, which is assigned by the Polish verb ‘ni lubia’ and marked by the regional Polish ending ‘-ów’ despite the fact that ‘bosonożki’ is a feminine noun and the expected Standard Russian form would be ‘bosonożek’.

While it can be concluded that in the second instance the Russian noun is morphologically integrated into the Polish base, just like we saw in Extract 38, its pronunciation

remains markedly Russian. Russian unstressed vowels differ from their stressed counterparts and are normally reduced or centralised (Avanesov 1956: 105–106; Bulanin 1970: 102), thus /a/ and /o/ are represented as [ʌ] or [ə]. The degree of centralisation depends on whether /a/ or /o/ are word-initial or not, and whether they are in the pre-tonic syllable. If /a/ is word-initial or pre-tonic, it tends to be pronounced as [ʌ], hence in Russian ‘bosonožki’ is normally pronounced as [bəsʌˈnoʃkʲɪ].

On the other hand, it is usually agreed that Polish has very little or no phonetic vowel reduction. Sawicka (1995) claims that vowel reduction of non-high vowels may occur only in phrase-final positions, as a result of the articulators slowing down and losing precision, but it is quite rare. Therefore, if ‘bosonožki’ were to be phonologically adapted into the Polish base, the expected pronunciation would likely resemble [busaˈnoʃki]. Speaker U’s pronunciation of the Russian noun is marked with a perceptible vowel reduction, characteristic of the Russian variant, which suggests that this lexical item is only integrated in terms of morphology, but not phonology, and should therefore be regarded as a prototypical single switch.

With regard to the above extract it is possible to speculate that the assignment of the masculine plural genitive case marker is influenced by the fact that ‘bosonožki’ is a plurale tantum noun, i.e. it does not have a singular form. It could therefore be argued that the speaker cannot tell whether it is a feminine or a masculine noun on the basis of the plural form alone, since in Standard Russian there is some overlap in the plural endings, e.g.: ‘noski’ [NOM masculine plural] – *socks*, ‘poloski’ [NOM feminine plural] – *stripes*. The most reliable way of establishing whether it is a masculine or a feminine noun is by looking at its singular form, i.e. ‘nosok’ [NOM masculine singular] – *a sock* ends with a consonant, while ‘poloska’ [NOM feminine singular] – *a stripe* ends with a vowel ‘-a’, characteristic of Russian feminine nouns. Since this method of verification does not apply to plurale tantum nouns, it could be assumed that the speaker automatically treats them as masculine and hence assigns

the Standard Polish masculine ending ‘-ów’. However, as Extract 41 demonstrates, in the regional variety of Polish spoken in Lithuania even those nouns which are clearly feminine are assigned the Standard Polish masculine ending:

Extract 41

- (1) Q: [...] u mnie w klasie było dużo _[RUS] **blondink-ów**
[...] at me in class was a lot of blondes
‘[...] There were a lot of blondes in my class.’

The lexical item ‘blondink-ów’ is a plural genitive form of the feminine noun ‘blondinka’. Just like in the previous example, when it is inserted into the Polish base it is assigned the regional Polish ending ‘-ów’, which suggests that it has been integrated as far as morphology is concerned. However, a closer analysis of its pronunciation reveals that it remains markedly Russian, and further confirms that we are dealing with a Russian switch, and not its Polish cognate ‘blondynka’. The most telling features are the reduction of /a/ in the pre-tonic syllable and the palatalisation of the phoneme /d/, which can be represented as [blɔn'dʲinkuf]. If it was the Polish lexical item, the phoneme /o/ would not be reduced and /d/ would be followed by /i/, and therefore it would not be palatalised. These observations suggest that ‘blondink-ów’ is only integrated in terms of morphology, but not phonology, and should therefore be regarded as a prototypical single switch.

One very significant finding emerges from the discussion of single Russian nouns employed in the Polish base. Apart from the local Polish accusative ending ‘-a’ and the genitive ending ‘-ów’, all the remaining inflections on switched nouns remain Russian. Recall from p.126 that case markers are classified as the ‘outsider’ late system morphemes since they are assigned by either verbs or prepositions. Following the System Morpheme Principle, outsider late system morphemes should always come from the matrix language. It can therefore be concluded that the Polish-Russian data display an obvious violation of one of the basic principles of the MLF model,

thus further confirming that the model does not hold universally and is particularly problematic to apply to such typologically closely related languages as Polish and Russian. The implications of this finding will be further addressed in the concluding chapter in section 7.2.

5.3.2 Verbs

Verbs that behave as single switches include infinitive and imperative forms, as well as present, past and future tense forms. They can be found in the discourse of every speaker with a total number of 680 tokens throughout the corpus. They tend to appear in their original form, i.e. they are not integrated by means of Polish inflectional endings and no phonological adjustments are made in order to integrate them into the receiving language. Examples of different forms of Russian verbs can be found in Extracts 42–45:

Extract 42

- (1) B: i z kim ona tam _[RUS] šljaetsja?
and with who she there goes out
'and who is she going out with [over] there?'

Speaker B employs 'šljaetsja' in the third person singular present tense form. The verb is imported into the base/matrix language without any morphological adjustments and can therefore be regarded as unintegrated. Moreover, 'šljaetsja' also appears to be unintegrated as far as its phonology is concerned. This is particularly evident in speaker B's pronunciation of the 'lja' sequence, where the Russian phoneme /l/ is clearly palatalised and can be represented as [lʲa]. In Standard Polish the phoneme /l/ is only palatalised before the vowel /i/, but never before any other vowel, thus if it were to follow the rules of Polish phonology the sequence would be realised as [la]. Moreover, the stress is retained on the first syllable, which is compatible with

the rules of Russian articulation.

Another example can be found in Extract 43:

Extract 43

- (1) G: on szybo _[RUS] **povzroslel**
he quickly grew up
'he grew up quickly'

The verb 'povzroslel' is used in the third person past tense form. Russian masculine singular verbs in the past tense do not take any person-endings after the word-final suffix. The same rule applies to Polish equivalent third person forms, which also end in /l/, e.g. 'powiedział' - [*he*] said. Therefore 'povzroslel' can be classified as a bare form, which does not require to be morphologically integrated into the Polish base. The pronunciation of the verb is kept distinctly Russian as well, with the stress falling on the last syllable, which suggests that the Russian lexical item patterns very much like a single switch.

Extract 44 contains one more example:

Extract 44

- (1) N: ja czekam póki Marta _[RUS] **vyšvyrnet**
I wait until Marta kicks out
'I am waiting until Marta kicks [me] out'

In the above extract the verb 'vyšvyrnet' is employed in the third person form. It can be classified as a present tense form in the perfective aspect, which is employed to express future meaning. The verb is imported from Russian with its original third person inflectional ending '-et'. Unlike the past tense form, where there is some overlap between the Russian endings and the Polish endings, the future tense forms in Polish are different from the Russian ones. It is therefore possible to conclude

with certainty that the Russian form ‘vyšvyrnet’ displays no overt morphological integration into the Polish base since the original Russian morpheme is preserved. The same can be said about phonological integration, for the verb is pronounced the same way it would normally be in Russian, with the stress on the first syllable.

Another example can be found in Extract 45:

Extract 45

- (1) Q: to wtedy jego zaczął _[RUS] **šipat’** to wtedy ((krzyczy)) ‘Jula (.) Jula’
 so then him began to pinch so then ((shouts)) Jula (.) Jula
 ‘so then [he] started to pinch him, so then [he] screams ‘Jula (.) Jula’

The above extract contains an example of the verb followed by the Russian infinitive. This type of pattern can be found across different languages and is very convenient for language alternation since the infinitive is not inflected for person, number or tense, and therefore requires no morphological integration in the base language. Moreover, the Russian verb ‘šipat’ displays no phonological integration into the receiving language. Its pronunciation and stress patterns differ from those of its Polish cognate ‘szczypać’. The Russian lexical item is pronounced as [ʃɛɪˈpatʲ], with the stress on the last syllable, whereas the Polish one as [ʃtʲɪpatɕ], with the stress on the penultimate syllable. Speaker Q’s variant closely resembles the Russian verb.

The forms of the verb discussed so far can all be classified as the ones that pattern and behave like single switches. They are imported into the Polish base without any overt morphological integration and are non-recurrent within the present corpus, i.e. they are used on only one occasion. As seen with nouns, the fact that switched verbs display no overt morphological integration does not mean that they violate the morpho-syntactic make-up of the sentence or render it ungrammatical. Due to the prototypical closeness between Polish and Russian the embedded Russian forms either coincide with their equivalent Polish forms, e.g. third person past tense ending [l] in

Extract 43, or are inserted in such slots where their presence does not violate the well-formedness of the rest of the sentence, e.g. the use of the imperfective infinitive in Extract 45.

As far as verbs that appear as single switches are concerned, examples analysed above demonstrate that they are transferred into the Polish base without any phonological adjustments, and most importantly, they retain their original ending. According to Myers-Scotton (2002: 263), endings of the verb are examples of late system morphemes since in subject-verb agreement it is the subject that assigns the endings. The fact that Russian late system morphemes can be found not only in embedded nouns, but also in verbs, further corroborates the finding that in closely related languages, such as Polish and Russian, original late system morphemes on switched items are not only permitted, but they are actually preferred. This in turn suggests that the System Morpheme Principle, stated on p.120, is actually false. Contrary to what the principle stipulates, the examples of switched nouns and verbs analysed in the present chapter show that late system morphemes need not always come from the matrix language.

5.3.3 Adjectives

The majority of Russian adjectives that appear in the data are forms that keep their original Russian ending, and thus display no overt morphological integration. There are 236 instances of the latter in the present corpus. On the other hand, there are 78 instances where Russian adjectives are integrated on the morphological, but not on the phonological level.

With a total of 236 instances, adjectives that appear in their original form are the most common in the present corpus. Some examples can be found in Extracts 46 and 47:

Extract 46

- (1) B: on taki jeszcze ^[RUS] **molodoj** mężczyzna
he such still young man
'he is still such a young man'

In the above extract the Russian masculine singular adjective 'molodoj' is employed in its original form. It displays no phonological integration, for it is stressed on the last syllable as it would normally be in Standard Russian.

Another example can be found in Extract 47:

Extract 47

- (1) A: tam polski sklep polskich sukienek ^[RUS] **vypusknyh** (.) ^[RUS] **nu** takich
there Polish shop Polish dresses prom (.) well such
(.) ^[RUS] **prazdničnyh**
(.) festive
'it's a Polish shop with Polish prom dresses (.) well, such festive [dresses]'

The above extract contains two examples of Russian adjectives, employed in the plural genitive form. In both instances, the adjectives are embedded into the Polish base in their original form, though they still fulfil the syntactic requirements of the given sentence since they are used in the genitive case. Neither adjective has been adapted on the phonological level since both of them are stressed in accordance with the rules of Russian pronunciation, i.e. 'vypusknyh' is stressed on the last syllable, whereas 'prazdničnyh' is stressed on the first syllable. This suggests that the three adjectives exemplified in Extracts 46 and 47 are instances of single switches.

The second most common group are adjectives that are morphologically integrated into the recipient language. Morphological integration pertains only to the ending, which is contracted to the same form that would normally appear on the Polish equivalent ending. Some examples can be found in Extracts 48–50:

Extract 48

- (1) D: _[RUS] **vo** renifery _[RUS] **mestne**
 here reindeer local
 ‘here are local reindeer’

The above extract contains an example of the Russian masculine plural adjective ‘mestne’. The Standard Russian masculine plural adjective ending ‘-yje’ has been contracted to ‘-e’ in line with Polish morphology, hence the adjective [‘m^ɨestnɨje] is pronounced as [‘m^ɨestne]. Although the adjective is integrated into the Polish base on the morphological level, it remains unintegrated as far its pronunciation is concerned. In Russian the initial consonant is normally palatalised as a result of the vowel /e/. If it were adapted according to the Polish phonological patterns, the adjective would likely be articulated as something similar to [‘m^ɨestne]. In speaker D’s version the initial consonant is clearly palatalised, which suggests that its pronunciation is more akin to Russian than Polish. It is therefore reasonable to assume that, in spite of the reduced suffix, ‘mestne’ is an example of a single switch.

Another example can be found in Extract 49:

Extract 49

- (1) J: to taka _[RUS] **protivna** scena
 this such repugnant scene
 ‘this is such a repugnant scene’

The above extract contains an example of the Russian feminine singular adjective ‘protivna’. The Standard Russian feminine singular adjective ending ‘-aja’ has been contracted to ‘-a’ in accordance with Polish morphology patterns. Similar to the previous extract, the adjective remains phonologically unintegrated. The Russian pronunciation of this adjective is characterised by the reduction of the unstressed vowel /o/, which could be transcribed as [pr^ɨʌ’tʲivn^ɨʌjə]. In Polish articulation the

vowel would remain unreduced and would be pronounced as [ɔ]. In speaker J's articulation the vowel is clearly articulated as the reduced [ʌ], which in turn suggests that the Russian adjective can be considered a single switch.

Extract 50 contains another example:

Extract 50

- (1) L: tam _[RUS] **tipa** byli _[RUS] **real'ne** sumy wiesz
there sort of were real sums you know
'there were sort of real sums, you know'

The above extract contains an example of the Russian feminine plural adjective 'real'ne'. In line with the previous examples, the Russian ending has been contracted to fit the patterns of the Polish morphology, hence '-ye' becomes '-e'. However, the adjective remains unintegrated on the phonological level. In Standard Russian the adjective 'real'ne' is pronounced as [rʲɪ'ɑlʲnʲɪ], whereas its Standard Polish cognate 'realny' is pronounced as [rɛ'alni]. The most notable difference is the palatalisation of /r/ in Russian, followed by the reduction of /e/ into [ɪ]. In speaker L's variant the consonant is articulated as palatalised and the vowel is reduced, hence the articulation of the adjective in this example is more akin to the patterns of Russian phonology. This suggests that the adjective can be regarded as a single switch.

5.3.4 Conclusions about Single Switches

The aim of the above discussion was to evaluate the presence of single switches in the present data in order to support or reject the Emerging Mixed Code Hypothesis, which states that intraclausal single switches and embedded language islands from Russian constitute the preferred type of language alternation, and are regularly and frequently inserted into the Polish base. While embedded language islands are discussed in the

next section, this section focused specifically on single switches, discussed under the lexical categories of nouns, verbs and adjectives.

One of the findings that emerge from the analysis presented above is that intraclausal switches, i.e. single switches together with embedded language islands, amount to 39% of all the Russian switches that occur in the present corpus, placing them right after discourse markers. While in total there are 2,605 instances of single switches, the three grammatical categories analysed in this section, namely nouns, verbs and adjectives, amount to 2,478 tokens and make up the most sizeable proportion of singly-occurring Russian lexical items. On the basis of this purely quantitative observation alone it can already be concluded that, after discourse markers, intraclausal switching involving single lexical items, constitutes the preferred type of language alternation in the present data, which supports one of the main premises of the Emerging Mixed Code Hypothesis.

This finding is further corroborated by the analysis of lexical items that appear as single switches. As demonstrated in the present chapter, in the great majority of instances they can be described as pertaining to random semantic categories and are non-recurrent. Moreover, they feature regularly in the discourse of all speakers in the present data, which suggests that they are wide spread and not limited to several speakers only.

The majority of lexical items exemplified in this chapter are inserted into the Polish base unintegrated. While some nouns and adjectives are assigned Polish case markers, they do not tend to be integrated as far as their pronunciation is concerned, which confirms that, despite some morphological adjustments, they pattern very much like ad-hoc insertions.

The purpose of the next section is to verify whether the validity of this premise can also be applied to constituents that consist of more than one lexical item, i.e. embedded language islands, which together with single switches are regarded as instances of

intraclausal switching of the insertional type and are said to be characteristic of the emerging mixed codes.

5.4 Embedded Language Islands

As stated in the introduction to this chapter, intraclausal switching of the insertional type includes not only single lexical items, but also larger constituents, which consist of more than one word and which are framed by the matrix language. Following Myers-Scotton (2005: 260–261), these larger constituents will be referred to as ‘embedded language islands’ (henceforth ELI). They can be defined as phrases or collocations within a bilingual clause, made up of words which show structural dependency relationships that make them well formed in the embedded language. This means that ELIs adhere to the morphosyntactic rules of the embedded language, and within an ELI, both content and system morphemes come from the embedded language. Moreover, they retain their original pronunciation. In other words, ELIs can be regarded as clusters of switched lexical items, which behave as a single unit inserted into the clause in the matrix language. In this respect they pattern very much like single switches and are therefore treated as instances of intraclausal switching of the insertional type.

Following Myers-Scotton’s definition, in order to be classified as an ELI, a collocation must fulfil several criteria. As mentioned earlier, the first criterion is that within an ELI all system and content morphemes come from the embedded language. The second important criteria is that the structure within an ELI is hierarchical, i.e. they show structural dependency relations, which means that some words are ‘higher’ than others in the structure in which they occur, e.g. the English phrase *glass of water* could act as an ELI since *glass* is the head noun while other elements are modifying it. In addition, some ELIs are also characterised by an additional feature, namely

the fact that ELIs are often collocations, i.e. they contain words that usually occur together. As Myers-Scotton (2005: 262) specifies, ELIs tend to be “formulaic in their composition”, which means that the words they include hardly vary at all. A good example of the latter would be set idiomatic units, such as the Russian idiom ‘na svoju golovu’ – *to one’s cost/misfortune*. According to Backus (2003), ELIs (or ‘units’ in his terminology) consisting of set collocations and idioms should be treated as single lexical entities, for in the process of their production they are accessed as single items, rather than being assembled together every time they occur. The label ELI will therefore be employed as a descriptive cover term to refer to multiword units embedded in the matrix language, including collocations and idioms.

After discourse markers and single switches, Russian ELIs constitute the third most numerous category of language alternation in the present corpus, with a total number of 904 tokens. Their prominent presence in the data can be regarded as a further piece of evidence in favour of the Emerging Mixed Code Hypothesis. In order to provide a more detailed overview of the types of ELIs that feature in the data, they will be subdivided into several smaller categories, namely noun phrases, prepositional phrases, and verb phrases. While the categories of noun and prepositional phrases are fairly straightforward, as will be demonstrated in the data analysis included in this section, the category of verbal phrases is more problematic. The problem lies in the fact that the verbs that appear inside verb phrases are often finite. According to some of the existing literature (e.g. Treffers-Daller 1991), the language of the finite verb dictates the base/matrix language. Following this view, it is difficult to decide whether the verbal phrase should be treated as an ELI or whether it should be treated as an instance of a Russian matrix language. This issue is addressed in greater detail in Section 5.5.3 and is brought up at this stage to indicate the shortcomings of the quantitative information presented here. Thus, the three most numerous categories consist of 185 tokens of noun phrases, 60 tokens of verb phrases and as many as 656 tokens of prepositional phrases. In addition, there are also 127 instances of so-

called ‘ragged’ multiword switches where the elements incorporated from Russian do not form any syntactic unit that would permit a formal syntactic account. They often resemble performance errors, showing the effects of limited planning time, and involve speakers using pauses, interruptions, hesitations, repairs and any other similar devices. While the analysis of ‘ragged’ switches could offer some insight into the cognitive workings of bilingual speakers and would therefore be of particular interest for psycholinguistic research, they will not be analysed in greater detail at this stage due to time and scope constraints, though they will certainly be identified as one of the potential areas for future research. The analysis presented in the rest of this section will focus on three major categories of ELIs consisting of noun phrases, verb phrases and prepositional phrases.

5.4.1 Noun Phrases

A noun phrase (henceforth NP) can be defined as a phrase that has a noun as its head element and which can function as a subject or an object. In the present data ELIs consisting of NPs all have adjectives, or adjective phrases, as their pre- or post-head attributes. In some data sets adjective and noun combinations are classified as internal ELIs, i.e. they are framed by the matrix language elements, e.g.: Cajun French/English ‘Ça c’est le *highest class* français’. - *That’s the French upper class.* (Brown 1986: 404, quoted in Myers-Scotton 1993: 152; translation from the original). The typically reported pattern is an NP consisting of a noun and an adjective, framed by the determiner from the matrix language. Unlike the examples reported in other studies, all the adjective and noun NPs that can be found in the present corpus are external ELIs. Overall, there are 185 tokens. Most of them consist of two lexical items, although there are also 33 instances of NPs that consist of two adjectives and a noun. There are no complex NPs in the present data, i.e. noun phrases that contain a clause, which may be either a relative clause or a clause in apposition. Some

examples can be found in Extracts 51 and 52:

Extract 51

A is talking about her business activities.

- (1) A: ja teraz woża z Polski ^[RUS] **koljaski detskie**
I now transport from Poland buggies children
'I now transport children buggies from Poland'

In the above example the ELI is an NP with a noun as its head and an adjective as its attribute. Both the noun and the modifying adjective are imported into the Polish base with their original accusative feminine plural endings and are pronounced exactly as they would normally be in Russian. The Russian phoneme /lʲa/ in the 'lja' sequence is clearly palatalised and can be represented as [lʲa]. Recall from the previous section that the phoneme /l/ is never palatalised before back vowels in Standard Polish, thus if it were to follow the rules of Polish phonology the sequence would be realised as [la]. Moreover, the unstressed phoneme /o/ in the pretonic syllable is articulated as it would normally be in Russian, i.e. as [ʌ]. Likewise, the initial phoneme /d/ is palatalised as a result of the vowel /e/ and the entire lexical item is pronounced as [dʲetskʲɨjə], which equals Russian pronunciation. The fact that the pronunciation and morphology remain unchanged confirms that the collocation exemplified above behaves like a typical ELI, embedded into the Polish base.

Another example can be found in Extract 52:

Extract 52

- (1) S: [...] Nowy Rok spotykał koło ^[RUS] **avstrijskoj granicy**
[...] new year met near Austrian border
'[I] celebrated New Year's [Eve] next to the Austrian border'

In the above extract the Russian ELI consists of an NP 'avstrijskoj granicy', which is the complement of a prepositional phrase, headed by 'koło'. The ELI is inserted

into the clause with Polish as the matrix language along with its original Russian morphology and pronunciation. Most notably, the phoneme /v/ is realised as [f] in the adjective ‘avstrijskoj’, which is compatible with the rules of Russian phonology. If it was the Polish variant, the expected realisation would be [w]. Likewise, the articulation of the noun ‘granicy’ remains markedly Russian, with the phoneme /n/ strongly palatalised and realised as [nʲ], which confirms that the collocation exemplified above behaves like a typical ELI, embedded into the Polish base.

5.4.2 Prepositional Phrases

In addition to NPs, prepositional phrases (henceforth PPs) can be described as another numerous category of ELIs that feature in the corpus, with a total of 656 tokens. A PP consists of a preposition, which functions as its head, and a nominal complement, which may be realised by a noun or a noun phrase, a pronoun, an adjective, an adverb, another PP or a clause. As far as the present data are concerned, the most common combination are ELIs consisting of a preposition and a noun or a nominal phrase.

When it comes to PPs in the present corpus, it should be noted that a great number of prepositions in Polish and Russian are identical in form, e.g. Russian ‘na’ – Polish ‘na’ – *on, at, to*, Russian ‘v’ – Polish ‘w’ – *in*. While they are not always identical in function, when their usage patterns coincide speakers have an opportunity to follow either the rules of Polish or Russian morphosyntax. According to the existing literature, words that are part of two languages can facilitate CS from one language to the other. Following Clyne (1967, 2003), this process is commonly referred to as ‘triggering’. He further specifies that bilingual homophones, such as the two prepositions mentioned above, are particularly good candidates for trigger words. While the examples that feature in the present corpus do not exclude that this could also apply with regard to the Polish-Russian data, there is no proof that the co-

occurrence between switches and trigger words is not a matter of coincidence. This issue would require further investigation, supported by statistical analysis, which would go beyond the scope of this thesis. Therefore at this stage PPs will be analysed as ELIs without further insight as to what could have triggered their insertion.

The preposition ‘na’ - *on, at, to* is one of the most common prepositions in the present data and features in numerous ELIs. It can be followed by a noun in the locative case to express location, or by a noun in the accusative case to express direction. Extracts 53 and 54 contain examples of ELIs consisting of PPs headed by ‘na’ and followed by a noun in the locative case:

Extract 53

- (1) S: pamiętam patrza pracował na tej (.)_[RUS] **na mojke**
 I remember watch worked on that (.) on car wash
 ‘I remember I see he worked at that (.) at the car wash’

The Polish equivalent of the Russian PP ‘na mojke’ is ‘w myjni’. If this ELI were to be adapted into the base/matrix language as far as the morphology is concerned, the consonant ‘-k-’ would change into ‘-c-’ in analogy with Polish, e.g. ‘chust-ka’ [NOM] – ‘w chust-ce’ [LOC] – *a scarf – in a scarf*. In the above extract the Russian PP is imported into the Polish base without any adaptations, which suggests that it can be classified as a switch.

Another example can be found in Extract 54:

Extract 54

- (1) D: my_[RUS] **na parome** plynęli
 we on ferry swam
 ‘We travelled by ferry’

‘Na parome’ is an example of another Russian ELI consisting of a PP headed by the preposition ‘na’ and followed by a noun in the locative case. It is embedded

into the Polish frame without any adjustments. The Polish equivalent would be ‘na promie’, though it would be grammatically incorrect since the required construction is ‘płynęliśmy promem’.

The Russian preposition ‘na’ can also occur in ELIs expressing motion or direction. An example of this construction can be found in Extract 55:

Extract 55

- (1) M: ty lepiej jutro [przyjdź]_[RUS] **na peregovor** (.) trzeba
 you better tomorrow [come] on chat (.) must
 porozmawiać
 talk
 ‘you better [come] for a chat tomorrow (.) [we] need to talk’

In the above example the Russian masculine noun ‘peregovor’ is employed in the accusative case. The fact that the Russian noun is stressed on the last syllable further confirms its unintegrated status, since in Polish the stress is fixed on the penultimate syllable. It can therefore be concluded that, just like in the previous examples, the Russian ELI consisting of a PP is inserted into the Polish base without any overt integration.

5.4.3 Verb Phrases

Another group of relatively frequently switched ELIs are verb phrases (henceforth VP), with a total number of 60 tokens across the corpus. A VP can be defined as a phrase that has a verb as its head. If it is a verb used with a transitive meaning, then it will also be followed by a pre- or post-head string (its direct object).

It should be noted at the very outset that, unlike the previous types of ELIs, ELIs consisting of VPs are problematic since their base/matrix language tends to be ambiguous. This can be illustrated with the following example:

Extract 56

- (1) B: mówili że jej mąż _[RUS] **kreditov nabral**
said that her husband loans took
‘[they] said that her husband took out loans’

The above example consists of a main clause and a subordinate clause. The main clause is expressed in Polish, while the subordinate clause is bilingual, i.e. it consists of one Polish NP (‘jej mąż’) and one Russian VP (‘kreditov nabral’), which makes it difficult to decide whether the matrix language is Polish or Russian, and whether it is the Polish NP or the Russian VP that should be classified as an ELI. According to the existing research, both interpretations are in fact possible. Doron (1983) and Joshi (1985), for example, define the base/matrix language on the basis of the language of the first word in a sentence, whereas for Klavans (1985) and Treffers-Daller (1991) it is the language of a finite verb that dictates the base/matrix language. Some researchers deny the existence of a base language altogether (e.g. Meeuwis & Blommaert 1994), while others admit that identifying a clear matrix language is particularly problematic in closely related languages, since such languages have many morpho-syntactic patterns in common (Clyne 1987: 760).

According to Myers-Scotton (2002: 68), who developed the very concept of the matrix language, it should not be treated as a concrete language, but as an abstract structure. She explicitly states that “the Matrix Language is an abstract frame. This means it does not include actual morphemes nor is it isomorphic with any fully fleshed-out linguistic variety. Instead, it includes specifications about slots and how they are to be filled, based on directions from lemmas in the mental lexicon (ibid).”

As stated in the introduction to this chapter, the MLF model, which aims to account for ‘classic’ CS, is based on two principles, the Morpheme Order Principle and the System Morpheme Principle, which are also supposed to help identify the matrix language (Myers-Scotton 2002: 244). Recall that the Morpheme Order Principle

states that:

In mixed constituents consisting of at least one Embedded Language word and any number of Matrix Language morphemes, surface word (and morpheme) order will be that of the Matrix Language.

The System Morpheme Principle states that:

In Matrix Language+Embedded Language Constituents, all system morphemes which have grammatical relations external to their head constituents will come from the Matrix Language, i.e. the so-called ‘late’ system morphemes under the 4-M Model.

However, it is unclear how they apply in identifying the matrix language in those instances where it is ambiguous, or in closely related languages such as Polish and Russian, for Myers-Scotton (2005: 261) states that “within an Embedded Language island, all the content and system morphemes come from the Embedded Language and follow other well-formedness requirements of the Embedded Language (e.g. word order). This means that, within the island itself, the Morpheme Order and the System Morpheme Principles of the MLF model do not apply.”

Bearing in mind that the domain of analysis is the clause, and not the sentence, the subordinate clause in the above extract can be labelled as ambiguous. Following the Morpheme Order Principle, the matrix language could be either Polish or Russian, since the surface word order and the morpheme order is exactly the same:

Polish	jej mąż kredytów nabral
Russian	ee muž kreditov nabral
CS	jej mąż kreditov nabral

Following the System Morpheme Principle, the outsider late system morphemes within the ELLs would not be taken into consideration, namely the genitive plural case marker on ‘kreditov’, since it is assigned by the finite verb ‘nabral’ and both of

these lexical items are within the same constituent. The only other outsider late system morpheme in the above example is the third person singular verb inflection on ‘nabral’, since its form depends on the subject, i.e. it must look for information outside the element with which it occurs and therefore outside its immediate phrase. In theory, the ending on the verb is Russian, which suggests that outsider late system morphemes within the subordinate clause come from Russian, and therefore it is Russian that should be regarded as the matrix language. However, if we assume that the speaker kept Polish as the matrix language during the transition from the main clause to the subordinate clause, and the verb was assigned a Polish outsider late system morpheme, it would still end in [l]. Although the pronunciation, and in particular the word-final stress on ‘nabral’ confirm that it is the Russian lexical item, the high degree of congruence between the two codes, as well as the fact that the ELI in question is a structurally indispensable constituent of the subordinate clause, make it difficult to state with certainty what the matrix language is in the above extract. It is therefore plausible to classify the above example as ambiguous as far as the matrix language is concerned.

It appears that the same problem occurs whenever the Russian ELI contains a finite verb, as can be further seen in Extract 57:

Extract 57

- (1) B: kolega ^[RUS] **pokazyval emu** brata swego ^[RUS] **na fotografijah** [...]
 kolega showed him brother own on photos
 ‘A friend showed him his brother on photographs.’

The above example contains several ELIs, though what they are depends on whether Polish or Russian is assumed to act as the base/matrix language. Assuming that Polish is the base language, one ELI would be a VP, ‘pokazyval emu’, while the other ELI would be a PP, ‘na fotografijah’. If we assume that Russian is the matrix language, then the above example would contain a single switch ‘kolega’ and an ELI

consisting of an NP ‘brata swego’. As with the previous example, the Morpheme Order Principle is not particularly useful in establishing the matrix language since the word and morpheme order for the above example is the same in Polish and in Russian:

Polish	kolega pokazywał mu brata swego na zdjęciach
Russian	drug pokazyval emu brata svoego na fotografijah
CS	kolega pokazyval emu brata swego na fotografijah

Following the System Morpheme Principle, the two outsider late system morphemes whose form is assigned by an element outside their immediate projection is the third person singular past tense inflection on ‘pokazyval’, which depends on the subject, and the genitive singular case marker on ‘brata’, assigned by the verb. Both of these morphemes are identical in Polish and Russian, which means that the choice of a Polish outsider late system morpheme or a Russian outsider late system morpheme would not violate the syntactic make-up of the sentence, which once again makes it difficult to decide which code should be regarded as the matrix language in the above example. Moreover, this observation can be taken as a further confirmation that it is not just the Morpheme Order Principle, but also the System Morpheme Principle that does not apply in such typologically closely related languages as Polish and Russian.

The observation that in some instances the base/matrix language is difficult to identify is supported by Auer and Muhamedova (2005), who analyse instances of language alternation between Kazakh and Russian, as well as Latin and Early New High German. They conclude that ELIs may be governed by the matrix language, but the matrix language itself may be influenced by the embedded language. Moreover, their analysis of bilingual subordinate clauses demonstrates that in some instances the lexical elements are so intertwined with each other that the matrix language is impossible to identify. The authors conclude that “often, there is no monolingual code which can be taken as the point of reference”, which means that the analysis of mixed

clauses should take “the syntactic structure of the mixed utterance as the starting point, rather than the monolingual ‘codes’ which these mixed utterances seem to refer to” (Auer and Muhamedova 2005: 52).

As far as the present data are concerned, it is important to acknowledge that there are instances where the base/matrix language cannot be unambiguously labelled as ‘Polish’ or ‘Russian’, but should be regarded as the composite of both, as in the two extracts quoted above. However, as demonstrated in numerous other examples analysed within this chapter, in the majority of cases it is possible to state that Polish is the base language, while Russian is the embedded language.

Following the above discussion, the only VPs that can be unambiguously labelled as ELIs inserted into the Polish base are the ones that are headed by the verb in the infinitive form. Some examples can be found in Extracts 58 and 59:

Extract 58

- (1) T: ty u siebie w maszynie umiesz _[RUS] **bak vydut’?**
 you at own in car know how tank blow out
 ‘Do you know how to blow out a tank in your car?’

In the above extract the clause is predominately Polish with a Russian ELI ‘bak vydut’ inserted at the very end. The ELI consists of a VP complement, headed by a verb in the infinitive form and its direct object in the pre-head position. As with the previous two examples, the Morpheme Order Principle is not particularly useful in confirming that Polish is the matrix language since the word and morpheme order for the sentence is the same in Polish and in Russian:

Polish	ty u siebie w samochodzie umiesz cysterne odpowietrzyc
Russian	ty u sebja v mašine umeeš’ bak vydut’
CS	ty u siebie w maszynie umiesz bak vydut’

Following the System Morpheme Principle, ‘umiesz’ is the finite verb, whereas the Russian ELI, consisting of the infinitive VP ‘bak vydut’ acts as its infinitival complement. The infinitive ending ‘-t’ is an outsider late system morpheme, contingent upon the finite verb. Assuming that ‘maszyna’ is a lexical borrowing, apart from the Russian ELI, all the content and system morphemes in the above clause originate from Polish, including the subject and the finite verb, which suggests that Polish also acts as the base/matrix language.

Another example can be found in Extract 59:

Extract 59

- (1) D: ja pojada _[RUS] s **Turkami napivat’sja**
 I will go with Turks to get drunk
 ‘I’ll go to get drunk with the Turks’

Once again, the Morpheme Order Principle is not particularly useful in establishing the matrix language since the word and morpheme order for the sentence is the same in Polish and in Russian:

Polish	ja pojadę z Turkami się upić
Russian	ja poedu s Turkami napivat’sja
CS	ja pojada s Turkami napivat’sja

A closer analysis of the above clause shows that the Russian ELI ‘s Turkami napivat’sja’ is a VP complement, headed by a verb in the infinitive form, and a PP (‘s Turkami’). As with the previous example, the infinitival ending in ‘napivat’sja’ is an outsider late system morpheme, assigned by the finite verb. However, the subject and the verb of the clause in which the ELI appears are both Polish, whereas the ELI is an adjunct, and can therefore be classified as an optional element of the secondary importance. Following this line of reasoning, Polish is more likely to be the base/matrix language in the above example, since the subject and the finite verb are both in Polish, whereas Russian is more likely to act as the embedded language.

5.4.4 Internal Embedded Language Islands

While the examples analysed in the previous section contain ELIs that are inserted into the Polish frame without any morphological integration, there are some examples in the data where ELIs do receive a matrix language element. Myers-Scotton (2005: 265) treats them as a sub-type of ELIs and refers to them as internal ELIs.

With an overall number of twelve tokens, it is safe to conclude that internal ELIs are very rare in the present corpus. In all instances they are PPs, consisting of a preposition and a noun. Some example can be found in Extracts 60–62 below:

Extract 60

- (1) S: za cieżko już [pracować]_[RUS] **na stroj**-ce
too hard already work on construction
‘[it is] too hard to work on construction’

The above example contrasts with the one in Extract 53, where the Russian ELI ‘na mojke’ is inserted into the Polish frame without any adaptations. Although it retains its distinct Russian pronunciation, the feminine noun ‘strojka’ is adapted to the patterns of the Polish morphology by changing the Russian consonant ‘-k-’ into the Polish ‘-c-’, in line with the Polish locative case ending. The resulting construction is an internal ELI framed by the matrix language element.

Extract 61 contains another example of a Russian internal ELI:

Extract 61

- (1) F: kto_[RUS] **na stank**-u umie
who on machine knows
‘who knows [how] to use the machine’

The above example is problematic since it can be interpreted in two ways. The first interpretation would be that the Russian masculine noun ‘stanok’ takes the Polish

masculine locative inflectional ending ‘-u’, as in ‘mrok’ [NOM] – ‘w mroku’ [LOC] – *in the darkness*. This would mean that ‘na stank-u’ is another instance of an internal ELI framed by the matrix language element. The second interpretation would be that the entire ELI comes from Russian, but instead of the expected locative ending ‘-e’, it is assigned the Russian locative ending ‘-u’, which can be found in such singular masculine items as ‘dym’ [NOM] – ‘v dymu’ [LOC] – *in smoke*, ‘škap’ [NOM] – ‘v škafu’ [LOC] – *in the wardrobe*. If phonological features were to be taken into consideration, then the latter option would seem more plausible, since the stress is kept on the last syllable, which is characteristic of Russian realisation, and not Polish.

Internal ELIs can also consist of ‘na’ followed by the accusative case to express motion or direction, as in Extract 62:

Extract 62

- (1) Q: tak a jutro mama pójdzi _[RUS] **na wystavka**?
 so and tomorrow mother will go on exhibition
 ‘so will mother go the exhibition tomorrow?’

Recall from Extracts 30 and 31 that the Standard Russian accusative ending for feminine nouns is ‘-y’, thus the expected ending of ‘vystavka’ would be ‘vystavku’. In the above extract the noun takes the local Polish ending ‘-a’ in the accusative form, which suggests that this is another instance of an ELI framed by the matrix language element.

5.4.5 Conclusions about ELIs

The aim of the above discussion was to evaluate the presence of embedded language islands in the present data in order to support or reject the emerging mixed code hypothesis, which predicts that intraclausal switching of the insertional type constitutes the preferred type of language alternation in the present data. Judging from

the quantitative information presented in this chapter it can be concluded that after discourse markers and single switches, ELIs is the third group of most frequently switched lexical items, with a total of 904 instances. Just like single switches, in the majority of instances ELIs tend to be inserted without any morphological integration into the base language. An exception to this rule are internal ELIs, which receive a matrix language element that frames them in the base language. However, with only twelve tokens across the corpus their presence cannot be considered as significant.

While NPs and PPs acting as ELIs are fairly straightforward, VPs appear to be more problematic since they often include a finite verb, which makes it difficult to decide whether the given VP should be treated as an ELI inserted into the Polish matrix language, or whether the VP itself is part of the Russian matrix language. As demonstrated in the discussion above, in some instances the matrix language will remain ambiguous, which is largely conditioned by the fact that Polish and Russian are two closely related languages. The only VPs that can be labelled as ELIs inserted into the Polish base are the ones that are headed by the verb in the infinitive form.

On the basis of the quantitative information, as well as the fact that ELIs pattern very much like single switches with regard to integration, it can be concluded that the second premise of the Emerging Mixed Code Hypothesis is also true as far as ELIs are concerned.

The aim of the following section is to test the third and final premise of the the Emerging Mixed Code Hypothesis, which predicts that there is evidence of convergence between the two codes on the grammatical level. The reason why grammatical convergence is included as part of the hypothesis is that it is assumed here that CM arises as a result of intense language contact and high levels of bilingualism, which leads to not only an increase in the density of switching on the lexical level,

but also affects the grammatical structure. In other words, evidence of prolonged language contact can be seen on the level of the lexicon and syntax. In order to test the grammatical effects of language contact, the analysis will focus on instances of structural borrowing, which include semantic borrowings and calques.

5.5 Structural Borrowing

As mentioned in the introduction to this chapter, structural borrowing refers to the process of copying of syntactic, morphological or semantic patterns (Haspelmath 2009: 39). As far as Polish-Russian language alternation is concerned, two types of structural borrowing can be singled out, namely semantic borrowings and calques. While not as common as intraclausal switches of the insertional type, structural borrowings, in particular calques, are quite productive in the present data and feature much more prominently than material borrowings (or loanwords), analysed earlier in this chapter. The fact that they are present in the corpus is quite significant, for it demonstrates that mixing between the regional variety of Polish and Russian goes deeper than the surface lexical level and can also be found in the underlying grammatical structure, which corroborates the third premises of the Emerging Mixed Code Hypothesis. The discussion presented below starts with examples of semantic borrowings, it is then followed with the analysis of calques, which may consist of a single lexical item (loan translation) or several constituents (structural borrowing), and it concludes with an overview of examples that display convergence on the abstract level, and CS on the surface level.

5.5.1 Semantic Borrowing

According to Wiebusch and Tadmor (2009: 590), in the process of semantic borrowing, also known as semantic extension, lexical items from the native language are assigned

the meaning of their equivalents in another language. In other words, a semantic pattern of the donor language is reproduced in the recipient language. Two remarks have to be made with regard to semantic borrowings in the present data. First, it should be noted that, with only four obvious examples, they do not feature nearly as prominently as other types of structural borrowing. Second, semantic borrowings, as well as calques, are notoriously difficult to identify, for they are disguised with the morphemes of the recipient language and can easily be taken for native language words. It should also be noted that the processes of semantic borrowing and calquing greatly resemble each other and it could even be argued that they should be regarded as one and the same phenomenon, especially since both of them involve copying of covert elements from the donor language. However, the more traditional distinction between the two types of structural borrowing will be maintained in the present study, thus semantic borrowings entail assigning a new shade of meaning to the existing native lexical item whose grammatical form remains unchanged. On the other hand, calques, often referred to as loan translations, entail not only a change in meaning but also a change in the morphological or syntactic structure in analogy with the donor language. Effectively, and importantly, they entail the creation of a new word in the target language, which semantic borrowing does not.

Some examples of semantic borrowings that feature in the present data can be found in Extracts 63–65:

Extract 63

Speaker C is commenting on the geese he saw in the picture.

- (1) C: oho (.) jakie zdrowe oni
 oho (.) how healthy they
 ‘oho! they are so robust’
- (2) D: _[RUS] nu (.) trzeba było zakroić jego
 yes (.) should was slaughter him

‘yes (.) [I] should have slaughtered it’

In Polish the literal meaning of the adjective ‘zdrowy’ – *healthy* is its only meaning. In Standard Russian, on the other hand, ‘zdorovyj’ can signify someone or something that is *healthy*, but it can also be employed in the colloquial speech with the meaning of *strong, fit, robust*. In the above example the speaker is commenting on the size of the geese, rather than their health, hence it can be assumed that the native Polish adjective ‘zdrowy’ is used in the sense of *fit, robust*, which has been borrowed from Russian.

Another example can be found in Extract 64:

Extract 64

- (1) K: chcesz być inny [RUS] **da?** (.) jakiś drugi?
 want to be different yes (.) some second?
 ‘you want to be different, yes? somehow different?’

In Standard Polish ‘drugi’ is usually employed as a numeral with the meaning of *second*. In the above extract it functions as an adjective and its meaning can be translated as *different*. The latter meaning has been borrowed from Russian, where a similar lexical item ‘drugoj’ can only be used as an adjective to denote something or someone who is *different, other, another*. The extension of meaning of the Polish cognate in analogy with the meaning of the Russian lexical item can be classified as another instance of semantic borrowing.

Extract 65 contains one more example:

Extract 65

- (1) L: chleb musi świątły
 bread probably bright
 ‘[this] bread is probably light’

The Polish adjective ‘światły’ originates from the noun ‘światło’ – *light* and is normally employed with the meaning of *bright*. A search in the National Corpus of Polish reveals a total of 118 entries, with a significant majority of collocations consisting of either ‘światły człowiek’ – *a bright man* or ‘światły umysł’ – *bright mind*. Its Russian cognate ‘svetlyj’ seems to have a broader range of meanings, which is further supported by the variety of collocations in which it can appear in the Russian National Corpus. Most notably, it can be used in the sense *light* to specify a shade of colour, e.g. ‘svetlye bezhevye štany’ – *light beige trousers*, ‘svetloe pivo’ – *light beer*. The equivalent form that could be used in the latter two examples in Standard Polish is the adjective ‘jasny’, but not ‘światły’. Therefore, the use of ‘światły’ in the above example can be regarded as another example of a semantic borrowing from Russian.

5.5.2 Calques

Calques are characterised by the copying of syntactic, morphological and semantic patterns from the donor language into the recipient language. Also known as ‘loan translations’, they can consist of either a single word or a fixed phrasal expression. Following Haspelmath (2009: 39), a calque can be defined as a complex lexical unit that was created by a morpheme-by-morpheme translation of the source unit. One example of a single word calque can be found in Extract 45 on p.149, where the seemingly Polish verb ‘naczał’ – *began*, employed in the third person singular form in the past tense, is calqued in analogy with the Russian form ‘načal’. The calqued lexical item looks very similar to the equivalent Polish form ‘zaczął’, except that the Polish prefix ‘za-’ has been replaced with the Russian equivalent form ‘na-’.

The process of calquing is very productive in the present data. While in the majority of instances calquing involves single lexical items, there are also numerous examples of syntactic calques, where a larger grammatical structure of a phrasal unit has been borrowed from Russian and filled with Polish lexical items (Silva-Corvalán 1995: 257).

The presence of both types of calques reveals additional information about the nature of language alternation that we are dealing with in the present data. According to Myers-Scotton (2002: 100), calques should be treated as examples of convergence, characteristic of composite CS, where the underlying abstract structure of the lexical item originates from the donor language but the surface level is that of the recipient language. It is therefore important to analyse the extent to which calques feature in the present corpus, as their presence can be regarded as another crucial piece of evidence in favour of the Emerging Mixed Code Hypothesis.

The first type of calques can be described as lexical items that consist of a single word. In the overwhelming majority of cases, this type of calquing affects verbs, which amount to 155 tokens. In contrast, only three instances of possible examples of nominal calques and sixteen instances of adjectival calques have been identified. Calquing can involve either only one morpheme, e.g. the prefix, or several morphemes, such as the prefix and the stem. An example of a noun used as a calque is discussed in Extracts 66, whereas further examples of verbal calques can be found in Extracts 67–69:

Extract 66

- (1) A: jak raz byli **wyjściowe**
how once were weekend
'as it happens, it was weekend'

The term for *weekend* that is commonly employed in Standard Polish is exactly the same as its English equivalent. The noun 'wyjściowe' is a morpheme-for-morpheme translation of the Russian 'vyxodnye'. The stem 'vyxod-' has been replaced by its Polish equivalent 'wyjści-', whereas the adjectival suffix and the masculine plural ending '-nye' by the Polish equivalent '-owe'.

However, it should be noted that in the above extract calquing goes beyond the single nominal form and the entire utterance appears to have been calqued from Russian.

The adverb ‘jak raz’ is a word-for-word translation of the Russian ‘kak raz’ for *as it happens*, the verb ‘byli’ is a local-Polish homophone of the Russian ‘byli’, and ‘wyjściowe’ is a morpheme-for-morpheme translation of ‘vyходnye’.

The example above raises an interesting psycholinguistic question of whether we are dealing with established lexical and structural units, already stored in the mental lexicon, or whether they are produced online. The range of lexical calques attested in the present corpus, some of which are analysed in greater detail in the remainder of this section, suggests that most likely these occurrences reflect the selection of already established units. Even though there is only one token of ‘wyjściowe’ in the present corpus, it is unlikely that this noun is an example of a first-time occurrence. The same observation can be applied to the adverb ‘jak raz’, which consists of two lexical items but conceptually represents a single unit. Its status of an established lexical unit, retrieved from the lexicon, can be further confirmed by referring to usage-based corpus data, since it appears on a total of thirteen occasions in the discourse of eight other speakers.

This question is even more intriguing when it comes to calques of the second type, namely structural calques, which are analysed in Extracts 72-74. They are characterised by the underlying abstract structure that clearly originates from Russian, but are filled with Polish lexical elements on the surface level. However, unlike single lexical calques, the status of structural calques cannot be verified by referring to usage-based corpus data alone, since the examples attested in the data are all single occurrences. It is possible that some of them do represent novel combinations, but without further psycholinguistic evidence it is impossible to tell with certainty which of them are composed productively, and which of them are already entrenched in the mental lexicon.

When it comes to calquing, it can be concluded that, unlike switching, this type of structural borrowing affects predominately verbs, which is corroborated by the fact

that there are 155 tokens of the latter throughout the corpus. Some of the examples of verbal calques consisting of a single lexical item are discussed in Extracts 67–69:

Extract 67

- (1) J: po prostu [mnie] **nadojadło** to teoretyczne gadanie
simply [me] fed up this theoretical talking
‘[I was] simply fed up with this theoretical talking.’

The seemingly Polish verb ‘nadojadło’, employed in the third person past tense form, is an example of a calque based on the Russian form ‘nadoelo’. There is no immediate cognate verb in Standard Polish, with the closest equivalent being the expression ‘mieć dosyć czegoś’ – *to have enough of something*. Not only has this lexical item been translated morpheme-by-morpheme into Polish, it has also been assigned Russian syntactic features. Just like its Russian equivalent, the calqued verb ‘nadojadło’ takes the long form of the indirect object ‘mnie’, expressed in the dative case. Its closest equivalent in Standard Polish would normally take the direct object, expressed in the genitive case.

Extracts 68 and 69 contain a few more examples of calques, which display even greater morphological changes:

Extract 68

- (1) U: nie **przytwarzaj się** ty tak (.) mów jak jest i [RUS] vse
not pretend you so (.) speak how is and all
‘do not pretend so [much] (.) speak as it is and that’s all’.

The lexical item ‘przytwarzaj się’, employed in the second person singular imperative form, is based on the Russian equivalent ‘pritvarjajsja’. What is interesting about this example is that calquing also affects the prefix and the stem. The Russian prefix ‘pri-’ is replaced by the Polish equivalent ‘przy-’, the stem ‘tvor-’ by the Polish stem

‘twarz-’, the imperative ending ‘-jaj’ by the Polish ‘-aj’ and the reflexive particle ‘-sja’ with ‘się’.

On the prosodic level, the stress stays on the last syllable of the verb, hence in both Polish and Russian variants the stress remains on ‘-jaj’/‘-aj’. However, this does not mean that ‘przytwarzaj się’ is not phonologically integrated into the Polish base. According to Karaś (2009), in the regional variety of Polish spoken in Lithuania it is common to stress the imperative forms on the last syllable.

A similar example can be found in Extract 69:

Extract 69

T is about to check his e-mail.

- (1) T: zaraz **przewierz**ym (.)_[RUS] devushki muszą pisać mnie
 right away check (.) girls must write me
 ‘we’ll check right away (.) girls must write to me’

The verb ‘przewierzym’ is used in the first person plural future tense form. It is based on the Russian form ‘proverim’, which has been translated into Polish morpheme by morpheme. The Russian prefix ‘pro-’ is replaced by the Polish ‘prze-’, while the stem ‘ver-’ by the Polish ‘wierz-’. The ending of the verb comes from the variety of Polish spoken in Lithuania, where the Standard Polish ‘-ymy’ is usually employed in its shortened version ‘-ym’. In the original Russian form and in the calqued verb the stress is on the penultimate syllable, which coincides with the rules of Polish prosody.

There are several instances of adjectives and adjectival participles that can be classified as calques, though with a total of sixteen instances they cannot be described as very common. An example can be found in Extract 70:

Extract 70

- (1) D: ja _[RUS] srazu myślał może jaki **pląsmasowy** czy co
 I right away thought maybe which plastic or what
 ‘I thought right away maybe it was [made of] plastic or something’

The adjective ‘pląsmasowy’ is based on the Russian noun ‘plastmassa’, which is pronounced as [pləs'masə]. Note that the phoneme /t/ in the stem is not pronounced, which suggests that the local Polish calque is based on the spoken variant rather than the written one. Hence, the stem ‘pląsmas-’ is borrowed from Russian via the spoken medium, and the Polish adjective ending ‘-owy’ is attached to it. In addition, the stress is moved from the second syllable to the penultimate syllable in order to conform to the patterns of Polish prosody.

An example of an adjectival participle can be found in Extract 71:

Extract 71

- (1) L: tam **zapuszczone** _[RUS] zdanie
 there neglected building
 ‘[there is] a neglected building there’

The adjectival participle ‘zapuszczone’ is a calque of the Russian equivalent form ‘zapušennoe’. The Russian stem ‘zapuš-’ has been borrowed into Polish, while the Russian ending ‘-ennoe’ has been replaced with the Polish equivalent ‘-one’. In addition, some changes have also been made on the phonological level. Most notably, the Russian [çç] has been replaced with the Polish [ʃtʃ] and the stress has been moved from the second syllable to the penultimate syllable.

The second type of calques that features in the present corpus are structural calques. They can be defined as multiple-word calques that reproduce a collocation from the source language with the lexical material from the receiving language. The resulting syntactic structures are not commonly used or even non-existent altogether in the recipient language. According to Myers-Scotton (2005: 271), structural calques

constitute examples of convergence, where all the surface-level forms come from one language, but with part of the abstract structure that underlies the surface-level patterns originates from another language. She further specifies that CS often precedes, or combines with, convergence. When the two processes co-exist, they give rise to ‘composite CS’ (referred to as CM in the present study). As Myers-Scotton explains (*ibid*), composite CS features surface-level morphemes from both languages (as in ‘classic’ CS), but also the abstract structure of the frame itself is a composite of two languages, even though one language remains as its main source.

While not as productive as single word calques, structural calques are fairly common in the present corpus, with an overall number of 83 occurrences. Instances of ‘pure’ convergence are slightly less common than instances of convergence with CS (or CM), with 36 instances of the former, and 47 of the latter. Both types of convergence will be discussed further in the remainder of this chapter.

Instances of ‘pure’ convergence are characterised by the underlying abstract structure that clearly originates from Russian, filled with Polish lexical elements on the surface level. Some examples can be found in Extracts 72–74:

Extract 72

A is talking about her friend who works as a nanny.

- (1) A: ona tam dziecko patrzy
 she there child watches
 ‘she’s looking after a child there’

The underlying structure of the phrase ‘dziecko patrzy’ originates from the colloquial Russian expression ‘smotrit rebenka’. This syntactic structure is not permitted in Polish with the meaning intended by the speaker and it could only be interpreted literally, i.e. *the child is looking [at something]*, rather than *the child is being looked after*. The intended meaning would be expressed with the Polish phrase ‘opiekować się dzieckiem’, which is syntactically different from the one employed in the example

above. The Russian structure is made up of a verb followed by the accusative case, whereas the Polish one consists of a verb followed by the instrumental case. This shows that speaker A borrows the underlying syntactic structure from Russian but uses surface-level lexical forms from Polish.

Another example can be found in Extract 73:

Extract 73

N is talking about his job interview.

- (1) N: potem za mną zaszedł drugi człowiek
then after me entered different person
'then a different person went in after me'

The above extract is a good example of multiple processes taking place at the same time. The underlying syntactic structure is based on the Russian equivalent construction 'potom za mnoj zašel drugoj čelovek', which differs from the corresponding Polish structure 'potem po mnie weszła inna osoba'. On the syntactic level, 'za mną zaszedł' is calqued on the basis of 'za mnoj zašel', where the preposition 'za' is followed by a pronoun in the instrumental case. This Russian syntactic construction is filled with corresponding Polish lexical items, i.e. the preposition 'za' and the pronoun 'mną' in the instrumental case. However, the process of calquing does not end here. The verb 'zaszedł' is calqued in analogy with the Russian 'zašel', where the Russian prefix 'za-' is employed instead of the Polish 'w-'.

In addition, the nominative phrase 'drugi człowiek' is calqued on the basis of the Russian equivalent form 'drugoj čelovek'. While both lexical items exist in Polish, none of them can be used with the meaning intended by speaker N in the above context. As illustrated earlier in Extract 64, in Standard Polish 'drugi' is employed as a numeral with the meaning of *second*, *another*. The extension of meaning of the Polish cognate in analogy with the meaning of the Russian lexical item can be classified as a case of semantic borrowing.

Similar conclusions can be made about the lexical item ‘człowiek’, which is employed in Standard Polish with the meaning of *man, human being*. While there is another lexical item that is used with the meaning of *person*, such distinction does not exist in Russian, where ‘čelovek’ is the only term that is employed to cover both meanings. The extension of meaning of the Polish cognate in analogy with the meaning of the corresponding Russian lexical item can be classified as another case of semantic borrowing.

One more example of a structural calque can be found in Extract 74:

Extract 74

N is talking about debit cards.

- (1) M: z kartecki nie można rozliczyć się
 from card not can pay
 ‘[one] cannot pay by card’

The syntactic construction exemplified above originates from the Russian equivalent ‘s kartočki nel’zja rassčitat’sja’. The verb ‘rassčitat’sja’ can take the instrumental case, e.g. ‘rassčitat’sja naličkoj/kartočkoj’ – *to pay by cash/card*, but can also be followed by the prepositional phrase ‘s kartočki’ – *by card*. A quick search in the Russian National Corpus confirms that both variants can be used. In Standard Polish, the equivalent verb ‘rozliczyć się’ would normally be followed by a noun in the instrumental case, e.g. ‘rozliczyć się gotówką/kartą’. This shows that the underlying syntactic structure is calqued in analogy with Russian, where such a construction is permitted and used on a regular basis.

In addition to the larger structural calque, the above extract also features the noun ‘karteczka’, which can be classified as an example of a single calque. The diminutive form ‘karteczka’ is based on the Russian equivalent ‘kartočka’ – *a credit or debit card*. In the present example the Polish cognate ‘karta’ has been supplied with the analogous suffix ‘-eczka’.

Finally, there are 47 instances in the corpus where convergence on the abstract level, in the form of structural calquing, combines with CS on the surface level. These examples are of particular interest, since they can be taken as a further piece of evidence in favour of the Emerging Mixed Code Hypothesis. Some examples involving nouns can be found in Extracts 75 and 76:

Extract 75

- (1) A: a tata pracuji _[RUS] **mehanikom**
 and father works mechanic
 ‘and father works as a mechanic’

The lexical item ‘mehanikom’ can be classified as a single switch. It is imported into the Polish base along with its Russian inflectional morpheme ‘-om’. The Polish equivalent form would be ‘mechanikiem’, with the stress falling on the penultimate syllable. The suffix, as well as the stress on the second syllable make it clear that the lexical form employed on this occasion is clearly Russian. However, there seems to be more going on here than just the use of a Russian lexical item. Although it is possible to come across a similar construction in Standard Polish, i.e. ‘pracuje mechanikiem’, the preferred and most commonly used variant would be ‘pracuje jako mechanik’ – *works as a mechanic*, with the noun expressed in the nominative case, followed by the conjunction ‘jako’. On the other hand, only one type of construction is permitted in Russian, namely ‘rabotaet’ – *works* + instrumental case to express the profession. It is therefore plausible to assume that the above example is an instance of word-by-word translation of the Russian construction, or, in other words, an instance of convergence on the abstract level and CS on the surface level, where one of the elements is taken from Polish while the other one is a switch from Russian.

Another example can be found in Extract 76:

Extract 76

- (1) D: a te wzrywają_[RUS] **dinamitom** (.) tam już ziemi nie ma
 and these explode with dynamite (.) there already soil not is
 ‘and these they explode with dynamite (.) there’s no more soil there’

The lexical item ‘dinamitom’ is imported into the Polish base along with its Russian inflectional morpheme. The Polish equivalent form would be ‘dynamitem’. While the two lexical items are clearly cognates, the item employed in the above extract retains its Russian pronunciation. The most immediately striking feature is the palatalisation of the phoneme /d/, which is followed by /i/ and is pronounced as [dʲ]. In the Polish cognate, the phoneme /d/ is followed by /i/, which does not cause palatalisation.

Similarly to the previous example, the underlying abstract structure appears to originate from Russian: ‘vzryvat’ dinamitom’ – *blow up with a dynamite*. Although its Polish equivalent is relatively similar, namely ‘wysadzić dynamitem’, there is evidence which suggests that the abstract structure in the above example originates from Russian. The first piece of evidence is the verb ‘wzrywają’, which can be described as a calque of the Russian form ‘vzryvajut’ – *[they] blow up*. Although ‘*wzrywać’ does not exist in Standard Polish, the prefix ‘wz-’ for ‘up’ and the morpheme ‘rywać’ for ‘tear’ can be found in other lexical items, which allows to create a similar lexical item in the regional variety of Polish. The second piece of evidence is the use of the Russian lexical item ‘dinamitom’, along with its original inflectional suffix ‘-om’. This suggests that, compared to the previous extract, Extract 76 contains even less adaptations on the surface level since only the verb is altered in order to fit the Polish frame. Therefore, the above example can also be classified as another instance of convergence with CS, characteristic of CM and emerging mixed codes.

Some examples including verbs can be found in Extracts 77 and 78:

Extract 77

(1) N: kiedy skończysz? ((laughter))
when finish?
'when will you finish?'

(2) M: _[RUS] **zatrudnjajus'** odpowiedzieć
find hard to reply
'[I] find it hard to reply'

Speaker M employs 'zatrudnjajus'' in the first person singular present tense form. The verb is imported into the base/matrix language without any morphological adjustments and can therefore be regarded as unintegrated. Moreover, 'zatrudnjajus'' *odpowiedzieć*' is another example of convergence with CS, where the underlying abstract structure is from Russian, whereas the surface level structure contains lexical elements from both languages. The abstract structure is the Russian set collocation 'zatrudnjajus' otvetit'', which does not have an immediate cognate in Polish, the closest equivalent being 'trudno mi odpowiedzieć' – lit. *it is difficult for me to answer*. Speaker M therefore employs the Russian underlying structure but keeps one of its elements in the original form and translates the second element.

Another example can be found in Extract 78:

Extract 78

(1) S: to wziął w jednej _[RUS] **uglubil** (.) z drugiej troche też skroił wziął
so took in one deepened (.) from second a bit also cut took
'so [he] took and made [a cut] deeper from one side (.) [he] took and cut a bit
off from the other side too'

The verb 'uglubil' is used in the third person past tense form. Russian masculine singular verbs in the past tense do not take any person-endings after the suffix '-l'. The same rule applies to third person equivalent forms in Polish, which also end in '-ł', e.g. 'powiedział'. Therefore 'uglubil' can be classified as a bare form, which does not require to be morphologically integrated into the Polish base. The pronunciation

of the verb is kept distinctly Russian as well, with the stress falling on the last syllable, which suggests that the Russian lexical item patterns very much like a single switch.

Moreover, the above extract contains another interesting example of structural calquing. The sequence ‘skroił wziął’ is calqued in analogy with the Russian ‘srezal vzjal’ – lit. *cut took*. While ‘wziął’ does exist in Polish, in the above context it would be redundant since it can only be employed in its literal meaning, i.e. to denote that someone lay hold of something with one’s hand, thus the sentence would normally end with ‘skroił’. It should also be noted that ‘skroił’ in itself is another example of calquing involving a single lexical item where the Russian prefix ‘s-’ is added to the Polish lexical item ‘kroił’ in analogy with ‘srezal’.

5.6 Conclusion

To summarise, the main aim of the present chapter was to assess the evidence in favour of the remaining two premises of the Emerging Mixed Code Hypothesis. The quantitative and qualitative information presented in the first half of this chapter can be taken as one piece of evidence in favour of the second premise of the hypothesis, which states that intraclausal switching, consisting of single switches and embedded language islands is the preferred type of language alternation in the present corpus. In line with the premise, one of the basic but crucial findings of the present chapter is the fact that the data contain a broad inventory of random lexical items from Russian that appear as intraclausal switches in a Polish base/matrix language. In total, they amount to 3,521 tokens and correspond to 39% of switches, which means that after discourse markers, they constitute the preferred type of language alternation among the respondents. In the majority of instances these lexical items are used on only one occasion, which means that they tend to be non-recurrent. Moreover, due to their

random nature and a great variety of forms involved it would be difficult to group them into separate semantic categories. This shows that Polish-Russian language alternation can be described as ‘advanced’ and that the speakers rarely depart from the switched mode of speaking.

The analysis of intraclausal switches presented in this chapter also reveals another important finding about Myers-Scotton’s MLF model, which was employed as the theoretical framework. Contrary to what the model predicts, late system morphemes from the embedded language are permitted in the matrix language, i.e. in a great majority of instances Russian lexical items imported into the Polish base tend to retain their original inflectional endings. The importance of this finding stems from the fact that the present data show that Myers-Scotton’s prediction does not necessarily apply to such typologically closely related languages as Polish and Russian. This finding is not exactly new, since some of the researchers, most notably Francescini (1998) and Berruto (2005), have already demonstrated that it is problematic to apply the MLF model when analysing CS between closely related languages or dialects of the same language. One possible explanation is that structural overlap is eased when a grammatical category from one language is fairly analogous to the corresponding category in the other language, thus bilinguals with typologically similar languages share representations for similar morpho-syntactic constructions and do not require to make as many adjustments in language contact situations as with typologically distant languages. This is also the reason why the MLF model works best when language alternation involves two typologically distant languages.

The third premise of the Emerging Mixed Code Hypothesis, which states that there is evidence of convergence of the regional variety of Polish towards Russian on the grammatical level, which manifests itself in the presence of structural and single-word calques, is also supported by the data analysis, presented in the second half of this chapter. Calques, which can consist of single lexical items, as well as larger syntactic

units are relatively prominent in the present corpus. Their presence can be regarded as significant, for it demonstrates that the degree of language mixing goes beyond occasional insertions on the surface level, but also affects the abstract structure of the frame itself. In addition to calques ‘proper’, the data also contain instances of the so-called ‘composite CS’, which involves convergence on the abstract level, as well as CS on the surface level. Following the MLF model, convergence with CS is characteristic of CM and emerging mixed codes, hence its presence in the present data can be regarded as another strong factor in favour of the hypothesis.

On the basis of the analysis presented in this chapter it can be concluded that there is enough supporting evidence to qualify the Polish-Russian language alternation as an instance of an emerging mixed code. However, as stated in the introduction, the presence of one type of language alternation does not exclude the other, hence the aim of the next chapter is to verify to what extent ‘classic’ or prototypical CS also features in the present corpus.

Chapter 6

Interclausal Switches

While the main focus of the previous chapter was on intraclausal switching of the insertional type within the clause, this chapter is concerned with a different type of switching, namely switching between clauses, which is characteristic of ‘classic’ or prototypical CS, as defined by Auer (1984). One of the defining features of interclausal switching is that a given clause is expressed in a different matrix language from the preceding and/or following one and tends to perform a discourse-related function. Since Polish is the base/matrix language of the great majority of clauses in the present data, interclausal switching occurs when it is juxtaposed with the clause that is expressed in Russian, thus the overall focus of the present chapter are clauses with Russian as the base/matrix language.

The analysis presented in this chapter pursues a twofold aim. First of all, it aims to assess the evidence in favour of the ‘classic’ CS hypothesis, which states that 1) there is a clear preference for one language of interaction at a time, 2) there is a change in the base language from the switching point onwards and the return to the previous language of interaction is not predictable, 3) it is possible to identify at least a tentative discourse-related function that CS performs and 4) there is no evidence of convergence on the grammatical level.

On the basis of the analysis presented in Chapter 5, it can already be concluded that the first premise of the ‘classic’ CS hypothesis is not borne out by the present data. Intraclausal switches of the insertional type feature very prominently in sentences with Polish as the base/matrix language, which suggests that the mixed mode of speaking is the preferred kind of interaction. In order to further test the validity of this conclusion, the data analysis presented in this chapter will focus on the quantitative information, i.e. how many tokens of unambiguously Russian clauses can be found in the data, which will allow us to assess the extent to which switching of the alternational type features in the present corpus. In order to test the second and the third premises of the hypothesis, the grammatical and pragmatic properties of Russian clauses will be analysed, i.e. whether they tend to be main or subordinate, complex or simple, and whether it is possible to identify at least a tentative discourse-related function that the clauses could potentially perform. While the fourth premise of the hypothesis has already proved to be incorrect by the data analysis presented in the previous chapter, the aim of this chapter is to test it even further by verifying whether there is any influence from Polish on the abstract and/or surface level in clauses with Russian as the base/matrix language.

As will be demonstrated in this chapter, only one of the premises of the ‘classic’ CS hypothesis is to some extent borne out by the data, namely that in some instances it is possible to identify at least a tentative discourse-related function that CS performs when two clauses are juxtaposed. This is particularly evident when CS is employed to set off direct and reported speech, e.g. in direct quotations. However, it should be noted that any claims about potential functions that CS performs are based on the subjective interpretation of the researcher, hence they should be treated as plausible guesses rather than empirically based. Moreover, it will be argued that there is no sufficient evidence in support of the remaining three premises. There is clearly no preference for one language of interaction since speakers very rarely depart from the mixed mode of speaking. Likewise, while ‘classic’ CS of the alternational type predicts

a change in the base language for a longer stretch of discourse, the present corpus demonstrates that the departure from the base language never goes beyond one or two turns, which further reinforces the view that speakers display a preference for language alternation of the insertional type. Finally, as demonstrated in the previous chapter, there are sufficient examples in the data to state that there is enough evidence of convergence of the regional variety of Polish towards Russian on the grammatical level. In addition, in this chapter it will be demonstrated that the convergence is unilateral, i.e. there is no evidence of convergence of clauses with Russian as the matrix language towards the regional variety of Polish.

The chapter starts with a brief overview of ‘classic’ CS and the discourse-related functions that it is said to perform. It is then followed by the analysis of specific examples from the corpus in order to evaluate to what extent Russian clauses feature in the data and whether their presence is significant enough to argue that this mode of switching is preferred over intraclausal switching of the insertional type. The analysis will include interclausal switches that occur within the same turn, as well as between turns. Finally, the discussion will focus on Russian switched clauses that feature Polish lexical elements.

6.1 Interclausal Switching and ‘Classic’ CS

Recall from the literature review, that in defining the ‘classic’ or prototypical CS within the Conversation Analysis framework, Auer (1984, 1999) distinguishes between CS of the alternational type and CS of the insertional type. In the former, there is a change in the base language from the switching point onwards and the return to the previous language of interaction is not predictable. In the latter, the required lexical item or several items are inserted into a surrounding passage in the other language, though the base language remains the same (Auer 2001: 445). As demonstrated

in the previous chapter, in the emerging mixed code intraclausal switching of the insertional type takes over and becomes the preferred type of language alternation. Though during the inception stage the two types of language alternation may co-exist, interclausal switching of the alternational type is nevertheless considered as one of the staple features associated with ‘classic’ or prototypical CS, described in the earliest accounts.

Since the main unit of analysis adopted for the purposes of the present study is the clause, interclausal switching can be defined as a type of language alternation where one clause is in one language and the next clause is in another language (Myers-Scotton 2002: 56). In other words, in interclausal switching a given clause is expressed in a different base/matrix language from that of the preceding and/or following one. According to Auer and Eastman (2010: 97–98), one of the main premises of the earliest studies of CS is that it occurs in a sociolinguistic context in which speakers orient towards one language at a time. Unlike intraclausal switching, made up of frequent insertions of one or several lexical items which do not carry any perceivable local meanings, interclausal switching of the alternational type is often discourse-related. According to Auer (1999: 313–314), by departing from the established language of interaction, participants wish to contextualise certain linguistic activities, which need to be interpreted in each individual instance. Since switching is activity-related, the syntactic units affected by language alternation must be large enough for such an activity to take place.

When analysing interclausal switching, it is important to differentiate between two aspects. From the discourse analysis perspective, interclausal switching may take place between turns, as well as within a turn. The analysis presented in this chapter will focus on interclausal switches that occur within a turn, and interclausal switches that occur between turns. From the purely syntactic perspective, interclausal switching within a turn can take place between two main clauses, or between a main clause and

a subordinate clause, which can be represented as follows:

CP1, CP2 = first and second clause; L1, L2 = first and second language

1. [CP₁ L1] [CP₂ L2]
[CP₁ L2] [CP₂ L1]
2. [CP₁ L1 [CP₂ L2]]
[CP₁ L2 [CP₂ L1]]

The above representation focuses not only on the fact that interclausal switching within a turn can take place between a main clause and a subordinate clause, but it also takes into account the direction of the switch. According to some researchers, the direction of the switch is not important when it comes to analysing discourse-related functions of CS since it is the juxtaposition between the two languages that contextualises different aspects of the conversation, and the same discourse function can be carried out by a switch in both directions (Alfonzetti 1998: 186; Milroy and Gordon 2008: 219). However, as Edwards and Dewaele (2007: 230) explain, paying attention to the direction of switching can give clues as to the dominant language of participants. With regard to the present data, in intraclausal switching the preferred direction of switching was clearly from the Polish base to Russian, which would suggest that Polish is regarded as the dominant language, while Russian as the subordinate language. It is therefore interesting to see whether the same observation holds true when it comes to switching involving larger syntactic units.

On the other hand, interclausal switching that takes place between turns can involve either one clause, or a complex sentence consisting of several clauses, which can be represented as follows:

T1, T2 = first and second turn; CP1, CP2 = first and second clause; L1, L2 = first and second language; () brackets indicate that the second clause is optional

1. T1: [CP₁ L1 ([CP₂ L1])]

T2: [_{CP1} L2 ([_{CP2} L2])]

2. T1: [_{CP1} L2 ([_{CP2} L2])]

T2: [_{CP1} L1 ([_{CP2} L1])]

As far as the present data are concerned, it should be said that with an overall number of 322 tokens, interclausal switching accounts for only 3% of all switches that feature in the present corpus and can be described as the least popular type of language alternation. Therefore, a point needs to be made as to why it is important to analyse this type of switching in the first place, as according to Myers-Scotton (2002: 56), interclausal switches are not of much interest for CS research and should be dismissed altogether. The question of inherent interest is a highly subjective matter and certainly not a sufficient factor to dismiss the whole subtype of language alternation altogether. Myers-Scotton justifies her view with the fact that in sentences with interclausal switches there is no opposition between the matrix language and the embedded language simply because they consist of two monolingual clauses, which means that configurations such as the ones represented above do not reveal anything about the way grammars of the two languages interact when they are in contact. While they may not contribute to our understanding of how grammars of the two languages are reconciled in CS the same way as intraclausal switches do, they may nevertheless reveal whether instances of language alternation perform any discourse-related functions, and thus facilitate making the distinction between CS and CM, which is directly linked with one of the main aims of the present study.

Before proceeding with the analysis of the actual examples from the present corpus, it is important to take a closer look at the discourse-related functions that interclausal switching is said to perform. They are briefly outlined in the following section.

6.2 Discourse-Related Functions of Interclausal

Switching

Switching for discourse-related purposes has been attested widely in many bilingual communities across the world (among many others: Auer 1984; Guerini 2006; Gafaranga 2007). As Auer and Eastman (2010: 98) neatly summarise, on the one hand, we can distinguish between discourse-related switches that are closely linked to what they refer to as “the machinery of conversational exchanges”, which includes turn-taking, quotations, repair work, sequence organisation, preference organisation, beginning or ending a story. On the other hand, there are discourse-related switches that may be described as ‘stylistic’, such as the contextualisation of a different modality (joking/serious, fake/true) and of interpersonal distance vs. intimacy, or expressivity vs. neutrality.

One of the earliest attempts to compile a taxonomy of discourse-related functions of CS was made by Gumperz (1982), who views CS as a ‘contextualization cue’, which can be defined as a “verbal or nonverbal cue that provides an interpretive framework for the referential content of a message” (1982: 131). The conversational functions that he singles out include the use of language alternation for 1) quotations or direct speech, 2) addressee specification, i.e. when the switch serves to direct the message to one of several possible addressees, 3) reiteration, i.e. when a message in one code is repeated in another code (literally or in a slightly modified form), usually to amplify or emphasise the message, and 4) message qualification, i.e. qualifying constructions that aim to add very specific additional information.

An even more detailed and exhaustive list of functions is proposed by Zentella (1997: 93–96), who analyses conversational functions of CS among Puerto-Rican children in New York. She employs Goffman’s (1979) concept of ‘footing’ as the underlying principle behind a broad variety of switches. As Goffman himself explains (1979: 5),

“a change in footing implies a change in the alignment we take up to ourselves and others present as expressed in the way we manage the production or reception of an utterance”. In Zentella’s data, the respondents used CS primarily to signal a change in footing via two approaches, namely the use of CS for realignment, whereby speakers create structure within their discourse, and the use of CS for appeal or control, whereby speakers manage their relationship with the interlocutor. The following eight realignment strategies can be singled out:

1. topic shift: the speaker marks a shift in topic with a shift in language;
2. direct and indirect quotations: the speaker recalls someone else’s speech and reports it directly or indirectly, not necessarily in the language used by the person quoted;
3. declarative/question shift: language alternation accompanies a shift into or out of a question;
4. future referent check and/or bracket: the speaker makes a side remark, marked by a shift in language, to make sure the listener knows his or her next referent;
5. seeking opinion or approval: the speaker seeks interlocutor’s approval, usually in the form of a tag;
6. role shift: switching occurs as the speaker shifts role from actor to narrator or vice-versa;
7. rhetorical ask and answer: the speaker asks a question and immediately answers it in a different language;
8. narrative frame break: the speaker departs from the narrative frame to evaluate some aspect of the story, to deliver the punch line, or ending.

As far as CS for appeal and/or control purposes goes, Zentella (1997: 95) singles out the following three strategies:

1. aggravating requests: where the switch intensifies/reinforces a command;
2. mitigating requests: where the switch is employed to soften the command;
3. attention attraction: where the change in language is employed to call for the listener's attention.

Apart from language alternation to signal a change in footing, Zentella (1997: 96) also identifies CS for clarification and/or emphasis. Within this function, she identifies some of the following strategies:

1. translations: the speaker switches for the translation of a statement, command, question etc.;
2. apposition and/or apposition bracket: CS marks the introduction of an appositional phrase that adds subject specification.

In Zentella's data, it was the latter category of switches that was the most frequent, followed closely by switching for realignment purposes. While switching for appeal and/or control purposes was not significant in Zentella's corpus, it was proven to be productive in some of the previous studies of Spanish-English CS, most notably Gumperz (1982) and Valdés (1981). As far as the present data are concerned, the most frequently occurring type of discourse-related interclausal switching falls under Zentella's realignment strategies. As will be demonstrated in the next section, speakers exploit the juxtaposition between Polish and Russian clauses in order to create structure in their discourse by setting off new verbal activities within the main flow of their conversation. The analysis presented below is divided into two parts, namely interclausal switching within a turn and interclausal switching between turns.

6.3 Interclausal Switching within a Turn

As stated earlier, overall there are 322 tokens of interclausal switches, which amounts to only 3% of all the switches attested in the present data. Out of 322 tokens, there are 217 tokens of interclausal switches within a turn, which may consist of switches between two main clauses, as well as switches between a main clause and a subordinate clause. With a total of 181 instances, the former definitely outnumber the latter, which amount to only 36 tokens. However, this observation is hardly surprising and can be explained by the fact that we are dealing with spoken language data, which are predisposed towards using less complex syntactic structures. Preference for employing two simple clauses, separated by a brief pause, rather than complex clauses, as reflected in the present data, is a good example of how speakers tend to limit syntactic complexity in spoken discourse. As Halliday (2006: 169) explains, “spoken language is more clausal (more and shorter clauses) whereas written language is more nominal (clauses longer and fewer).” With regard to the present corpus, the preference for simple clauses can be attributed to the fact that the data analysed here come from dialogues, which consist of a lot of very short turns and very few instances of longer monologic turns.

As far as the direction of switching within a turn is concerned, in the majority of instances the switch proceeds from Polish to Russian. For interclausal switches that take place between two main clauses there are 157 tokens where the first clause is in Polish, while the second clause is in Russian, and 24 tokens where the first clause is in Russian, while the second clause is in Polish. For interclausal switches that take place between a main and a subordinate clause there are 26 tokens where the main clause is in Polish, while the subordinate clause is in Russian, and ten tokens with Russian as the main clause and Polish as the subordinate clause.

Bearing in mind the analysis presented in the previous chapter, which dealt with Russian single switches and ELIs inserted into the Polish base, this tendency is not

particularly surprising. As stated earlier, the direction of the switch is, to some extent, indicative of the speaker's dominant language. The fact that in the present corpus switching tends to happen from Polish into Russian can be regarded as a partial confirmation of the general practice to maintain Polish as the base/matrix language, and Russian as the embedded language, though as demonstrated in the previous chapter, in some instances the base/matrix language is ambiguous, while in other instances Polish is the base language only on the surface level, whereas the abstract structure is a composite of both languages.

6.3.1 Switching between Main Clauses

With a total of 181 tokens, switches between two main clauses constitute the most numerous category of interclausal language alternation within a turn. There are 157 instances of when a switch proceeds from Polish to Russian, and only 24 instances of switching from Russian to Polish. Examples of both types are analysed in greater detail below.

As far as discourse functions of interclausal switches within a turn are concerned, it is reasonable to assume that they are closely related to the contrast created by the juxtaposition of the two languages. In the overwhelming majority of cases, speakers use CS to create turn-internal structure, or in Zentella's (1997) terminology, to signal a change of footing via realignment. The most common realignment strategies that occur in the present corpus are linked with the content and structure of the speaker's narrative, namely the use of CS to differentiate between direct speech and the speaker's own voice, to set off side remarks and departures from the narrative frame, to make comments and evaluations about what is being said, or to signal a change of topic. Some examples can be found in Extracts 79-82:

Extract 79

- (1) N: ona mówi _[RUS] **ja vas ponjat' ne mogu**
 she says I you understand not can
 'she says: I cannot understand you'

The extract presented above is an example of an interclausal switch where two clauses are used side by side, the first one has Polish as its base/matrix language, while the second one has Russian, hence the direction of the switch is from Polish to Russian. The speaker explicitly flags the direct speech with 'ona mówi' – *she says*, thus informing the listener that a direct quotation is about to be introduced. The use of Polish for the introductory part, and a switch into Russian for the direct quotation allow the speaker to create a contrast between the two parts of their narrative, as well as achieve authenticity by preserving the language used in the original utterance. However, as the existing research suggests, this may also be the case of a so-called pseudoquotation, i.e. one that is not formulated in the original language. Such use of CS can be described as 'iconic' because the change of language correlates with a difference in relations between the original speaker and content (Karrebæk 2003: 414).

Another example can be found in Extract 80:

Extract 80

- (1) V: staruszka przychodzi (.) _[RUS] **dajte mne dva talončika**
 old woman comes (.) give me two tickets
 'An old woman comes (.) give me two tickets'

The above extract is another example of interclausal CS involving two independent clauses, where language alternation proceeds from Polish to Russian. Likewise, switching helps the speaker to separate the main narrative from the direct speech, though unlike the previous example, the latter is not introduced explicitly. Instead, it is the alternation between Polish for the narrative part, and Russian for direct quotation that indicates the transition from one type of discourse to another.

A different kind of discourse-related function is exemplified in Extract 81:

Extract 81

- (1) A: dzwoni do mnie (0.1) Wika mnie źle (.)_[RUS] **a ja że**
phones to me (0.1) Wika me unwell (.) and I *emph. particle*
vrač zaočnýj
doctor extramural
'[she] phones me (0.1) Wika, I'm unwell (.) and I am the extramural doctor'

The above extract consists of three separate clauses, separated from each other by a pause. The switch from Polish to Russian takes place between the second and the third clause. The first clause frames the speaker's narrative, the second one contains a direct quotation of one of the characters involved in the story, while the third clause expresses a side remark made by the narrator. Unlike the previous two extracts, both the speaker's narrative and the direct quotation are expressed in Polish. The juxtaposition between Polish and Russian coincides with setting off the side remark from the rest of the narrative, though one cannot tell if the speaker did this deliberately. In fact, it should be noted that with the exception of direct speech, it is impossible to state with certainty whether language alternation coincides with the speaker's desire to contextualise a given discourse-related function, hence any interpretations included in the present data analysis should be treated as plausible guesses, rather than empirically based.

Another example can be found in Extract 82:

Extract 82

- (1) A: ona jego wcale nie wychowywała (0.2)_[RUS] **est' čego ispravljat'**
she him at all not raised (0.2) is what correct
'she didn't raise him at all (0.2) there is something to be corrected'

In the above extract the speaker switches from Polish, employed for the narrative part, into Russian in order to convey the subjective remark about what is being said.

It can be speculated that the contrast created by CS is used to highlight a break in the narrative frame, and thus indicate to the listener that the speaker is momentarily departing from her role as the narrator to that of the commentator. However, just like with previous examples, we do not know for sure if a change in the base language was deliberate since there is no empirical evidence available that would support or reject this observation.

While in Extracts 79–82 it is possible to make plausible guesses as to what could be the discourse-related function created as a result of the switch, there are also instances in the corpus where this is not the case. As far as switching from Polish to Russian is concerned, there are 148 tokens where some sort of discourse-related or stylistic function can be attributed, even if it is only tentative and not empirically based. The remaining 33 tokens are more ambiguous. Some of the examples of the latter can be found in Extracts 83 and 84:

Extract 83

- (1) D: następny latem do Turcji pojedzie (.) tam odpoczniesz (.) _[RUS] s
 next summer to Turkey will go (.) there will rest (.) with
russkimi pop'eś' normal'no
 Russians will drink properly
 'Next summer I'll go to Turkey (.) there [you can] rest (.) [you] can drink
 properly with the Russians'

In the above extract it is difficult to pin-point the exact discourse-related function or identify the stylistic effect created by language alternation between Polish and Russian. The switch is situated within the same, relatively neutral, narrative mode as the previous two clauses, thus there seems to be no apparent motivation that would encourage the speaker to switch into a different base/matrix language.

The same observation applies to Extract 84:

Extract 84

- (1) Q: tam w kiosku z kwiataami pracuji baba taka (.) [RUS] **ona mne ne**
there in kiosk with flowers works woman such (.) she me not
nnavitsja
like

‘There is a woman who works in the flower kiosk (.) I don’t like her’

Similar to the previous extract, the narrative mode in the Polish turn and in the Russian turn remains the same, hence it is difficult to assign a discourse-related or stylistic function to the switched clause. The only possible explanation would be that CS is employed to perform an expressive function in order to help the speaker convey a strong, in this case negative, emotion. The expressive function of CS has been thoroughly studied by Guerini (2006), who focused on Ghanaian immigrants in Italy. She found that her respondents usually employed Italian to convey positive feelings, such as amazement, delight or satisfaction for what one has just been told, whereas Akan was often, though not exclusively, employed for the expression of negative feelings (Guerini 2006: 128–133). It is possible to assume that CS serves a similar function in the above example, where the speaker switches to Russian to convey dislike of the woman working in the flower kiosk. However, such an interpretation is based solely on speculation and cannot be supported with further empirical evidence, hence it would be more reasonable to classify this example as another instance where CS does not appear to perform any immediately identifiable functions.

Finally, it is worth briefly focusing on those examples where switching involves two separate clauses, but proceeds from Russian to Polish. As noted earlier, examples of the latter are not very common, with an overall number of 24 tokens. However, what is interesting about them is the fact that in 21 instances out of 24 it is difficult to identify the exact function that CS performs. In the remaining three examples the switch from Russian to Polish seems to be associated with a change in topic. This observation contrasts sharply with what has been said about interclausal switches from Polish to Russian, where the exact opposite is true. This suggests that the

direction of the switch is associated not only with the speaker's dominance in one language or the other, as suggested by previous research, but it also determines the function that the switch is likely to perform. An example with an identifiable discourse function can be found in Extract 85, whereas Extracts 86 and 87 contain instances where the discourse function is less apparent:

Extract 85

R is criticising Q's girlfriend.

- (1) R: [...] słońce moje tam jeszcze coś (.) a jak trzeba coś
sun my there more something (.) but when must something
zrobić to jej nie ma
to do then her not is
'My sun and what not (.) but when something needs to be done then she's
not [there]'
- (2) Q: _[RUS] **uspokojsja ty eto ne to** (0.4) jutro bendzi deszcz
calm down you this not it (0.4) tomorrow will be rain
'Calm down, it's not that (0.4) Tomorrow it is going to rain.'

In the above example the interclausal switch takes place in turn (2), where the transition from Russian as the base/matrix language for the first clause to Polish as the base/matrix language for the second clause coincides with a change of topic. The Russian clause is employed to give an answer to speaker R's remarks, whereas the Polish clause, which occurs after a brief pause, introduces the weather as the new topic of conversation.

However, as Extracts 86 and 87 demonstrate, the discourse-related function is not always immediately identifiable:

Extract 86

- (1) T: chcesz skrzynka sprawdzić?
want mailbox check

‘Do you want to check your mailbox?’

- (2) S: _[RUS] **zahodi lučše v svoju** (.) popatrzaj co tam robi się
enter better in yours (.) look what there happening
‘Better check yours (.) look at what’s happening in there’

The first clause in turn (2) is expressed in Russian, whereas the second clause has Polish as its base/matrix language. The switch from Russian to Polish does not seem to coincide with any identifiable discourse function, as the narrative mode, as well as the overall mood of the turn remain the same, which suggests that the contrast between the two languages is not employed to create any stylistic effects either. Moreover, there is no reason to believe that the switch into Russian for the first clause could have been influenced by the previous turn, as it has Polish as the base/matrix language.

Another example can be found in Extract 87:

Extract 87

- (1) U: dobry kierowca nie zobaczył że _[RUS] **remen’** rwie się?
good driver not saw that belt tear itself
‘A good driver didn’t see that the belt is tearing apart?’
- (2) V: _[RUS] **prosto remen’ vyšel iz stroja i vse** (.) tak samo jak w
simply belt came out of order and all (.) same as in
samochodach często bywa
cars often is
‘The belt simply came out of order and that’s all (.) same as often happens in cars’

Similar to the previous turn, the contrast created by a switch from Russian into Polish in turn (2) does not seem to perform any discourse-related or stylistic functions since the narrative mode remains the same. While it could be argued that the switch into Russian in the beginning of turn (2) was triggered by the presence of a single switch

‘remen’ in the previous turn, or the use of the Russian discourse marker ‘prosto’, this interpretation seems to be highly unlikely. It has been demonstrated with ample examples in Chapter 4 that Russian discourse markers are regarded as ‘fused’ into the Polish-Russian mixed code and do not tend to trigger further switching. Likewise, the extracts analysed in Chapter 5 show that intraclausal switching of the insertional type is a widely attested type of language alternation in the present data, which features with great frequency and regularity in the discourse of each and every speaker. This suggests that there is no plausible stylistic, grammatical or discourse-related explanations for the interclausal switch in turn (2).

6.3.2 Switches between a Main and Subordinate Clause

As stated in the introductory part of this chapter, interclausal switches within a turn can also consist of a main clause in one language, and a subordinate clause in another language, though with a total of 36 tokens, they are less common than switches consisting of two independent clauses. As far as the direction of switching is concerned, the preference is still the same, i.e. in 26 tokens the switch proceeds from Polish to Russian, and from Russian to Polish in the remaining ten tokens. While in the majority of switches consisting of two independent clauses it is possible to identify at least a tentative discourse-related function, when it comes to interclausal switches between a main and a subordinate clause the function remains ambiguous in 32 tokens out of 36 (with switching in both directions). In the remaining four tokens, where the switch proceeds from Polish to Russian, the discourse-related function can be identified as setting off the reported speech. Some examples can be found in Extracts 88, 89 and 90:

Extract 88

- (1) K: coś mówił że on/ _[RUS] **on nedavno načal uglubljatsja v**
 something said that he/ he recently started to delve into

anarhizm

anarchism

‘[he] was saying something that he recently started delving into anarchism’

In the above example the main clause is in Polish, whereas the embedded clause is expressed in Russian. Interestingly, the conjunction ‘że’ is also in Polish, which, as the next two extracts will confirm, is the prevailing tendency in the present corpus, for there are no examples in the data where the embedded clause starts with a Russian conjunction. The pronoun ‘on’, which immediately follows the conjunction is treated as ambiguous and could originate either from Polish or from Russian, hence both variants are given in the transcription. Of course the phonetic analysis could potentially help to decide whether the pronunciation of this lexical item is more akin to Polish or to Russian. However, it would be difficult to arrive at a reasonable conclusion in this particular instance since the lexical item is very short and the only telling feature would be the pronunciation of the phoneme /o/, which tends to be articulated in a similar manner in the regional variety of Polish and Russian.

The reason behind the interclausal switch exemplified above could be attributed to the speaker’s intention to separate the main narrative from the reported speech, since Russian is employed to convey what has been said by someone else. However, in the present corpus examples of this type are very rare, and as Extracts 89 and 90 demonstrate, the motivation for switching tends to remain ambiguous:

Extract 89

- (1) T: do garażu ida płacząc (.) ni mogą patrzeć na swoją maszynę
to garage go crying (.) not can look at one’s own car
‘[I] go to the garage crying (.) [I] cannot look at my car’
- (2) S: ja myślał że ty/ [RUS] ty **svoe** Gol’fa **ljubiš**
I thought that you/ you your Golf like
‘I thought that you like your Golf’

- (3) T: lubia (.) lubia (.) ale kiedy on mnie lubić zaczn
 like (.) like (.) but when he me to like start
 ‘[I] like it (.) [I] like it (.) but when will it start liking me?’

Similar to the previous extract, the main clause is expressed in Polish, whereas the embedded clause is in Russian. The conjunction ‘że’ remains in Polish, whereas the pronoun ‘ty/ty’ is ambiguous, for it is exactly the same in both languages. Just like the pronoun ‘on’ encountered earlier, ‘ty’ consists of two phonemes and the only telling feature would be the pronunciation of /i/, which tends to be articulated in a very similar manner in the regional variety of Polish and Russian, and therefore cannot be employed as a reliable criterion to differentiate between the two phonetic systems.

Like with the majority of switches between a main clause and a subordinate clause analysed in this section, it is difficult to pin-point the exact function that CS may perform or to identify the reason why the speaker would decide to change the matrix language from one clause to another in the given example. From a pragmatic point of view, the discourse mode remains the same throughout the whole utterance, i.e. the switched segment does not perform any identifiable discourse structuring or stylistic functions. The only tentative explanation could be that the speaker treats the embedded clause as reported speech. However, given that he is conveying his own beliefs, rather than actual words, this interpretation seems to be unlikely and too speculative to be taken as solid empirical evidence.

Finally, Extract 90 contains an example of where the switch proceeds in the opposite direction, i.e. from Russian to Polish:

Extract 90

- (1) A: a ona ma męża [RUS] **ital’janca?**
 and she has husband italian
 ‘and her husband is Italian?’

- (2) B: _[RUS] **da** (.) już pięć lat jak oni _[RUS] **v brake**
 yes (.) already five years how they in marriage
 ‘yes (.) they’ve been already married for five years’
- (3) A: i jak ona z nim gada?
 and how she with him talks
 ‘and how does she talk with him?’
- (4) B: _[RUS] **vyučila ital’janskij jazyk** żeby z nim rozmawiać
 learned italian language so that with him speak
 ‘[she] learned the Italian language so that she could speak with him’

Like in the previous example, CS does not seem to perform any immediately identifiable discourse-related or stylistic functions in the extract exemplified above, even though the switch is in the opposite direction. The base/matrix language of the main clause is Russian, whereas the embedded clause, as well as the subordinating conjunction ‘żeby’ are expressed in Polish. This observation further corroborates the fact that in the present data there is a clear tendency to employ Polish subordinating conjunctions regardless of the directionality of the switch.

The analysis presented in this section focused on interclausal switches within a turn, which can consist of two independent clauses, as well as a main clause and a subordinate clause. Two interesting findings emerge from the above discussion. The first one is that in the majority of instances the switch is from Polish to Russian. This could be regarded as one piece of evidence that, despite being fully bilingual and proficient in both languages, speakers prefer to employ Polish as the base/matrix language of their interaction. The second interesting finding is that in most instances it is possible to speculate about the possible discourse-related function of language alternation. While this can neither be stated with certainty, nor further verified with empirical evidence, it leaves us with the possibility that speakers might exploit the juxtaposition between the two languages to structure their utterances and to contextualise certain verbal activities, but they do not necessarily switch with a particular discourse or stylistic

effect in mind. This is particularly true when it comes to switching from Russian to Polish in two independent clauses, as well as in switching in both directions between a main clause and a subordinate clause. One possible explanation is the fact that back and forth language alternation between Polish and Russian is the accepted mode of speaking in the speech community studied in this thesis and applies not only to intraclausal switches of the insertional type, but also to some instances of interclausal switches of the alternational type.

The main aim of the next section is to analyse interclausal switches which occur between turns and to verify whether any of the findings of the present section apply to switches that spread over several turns.

6.4 Interclausal Switches between Turns

The main characteristic of interclausal switches between turns is that the base/matrix language of the entire turn is Russian, whereas Polish is the base/matrix language of the previous and following turns. From the syntactic point of view, the clause itself may be either simple or complex, or, if it constitutes a longer monologic turn, it may even be a clause complex, which can be defined as “a sequence of structurally related ranking clauses” (Halliday 2006: 169). In order to be qualified as an interclausal switch, a given turn must consist of at least two lexical items that show structural dependency relationships, e.g. a noun and a verb. Turns that consist of only one lexical item, such as an agreement particle, or a succession of two discourse markers were quantified as instances of single switches, even though they constitute a separate turn. Following this criteria, the total number of interclausal switches between turns that appear in the present corpus amounts to 105 tokens. Apart from several ambiguous instances, Polish is by default employed as the base/matrix language, whereas clauses with Russian as the base/matrix language, which are the central focus

of the present chapter, can be regarded as a departure from the norm. Therefore, when it comes to language alternation between turns, the direction of switching is irrelevant, since it is always from Polish to Russian and then back to Polish. By contrast, what is relevant is the return to the original base/matrix language, i.e. at what point the speaker ends the switch into Russian and reverts back to Polish. According to the second premise of the ‘classic’ CS hypothesis, the return to the original base language from the switching point onwards is not predictable. This suggests that if ‘classic’ CS were the preferred type of language alternation in the present data then we would expect to find longer stretches of discourse, encompassing several turns, where speakers maintain Russian as the base/matrix language. However, it should be stated straight away that this premise is not borne out by the present data since utterances with Russian as the base/matrix language never go beyond a single turn. This finding further confirms that language alternation of the insertional type, characteristic of CM, is the preferred mode of switching in the present data. Moreover, it demonstrates that the first and second premises of the ‘classic’ CS hypothesis are false.

In almost half of the instances of switching between turns, i.e. 49 tokens, it is possible to assign at least a tentative discourse-related or stylistic function performed by the switched turn. Similarly to switches between two independent clauses that occur within a turn, CS functions as a realignment strategy and may be employed to differentiate between direct speech and the speaker’s own voice, to set off side remarks and clarifications related to the main narrative, or to make subjective comments or evaluations about what is being said. In addition, switches between turns can also perform what Auer and Eastman (2010: 98) describe as ‘stylistic’ function, namely contextualisation of a different modality, particularly a transition from a neutral mode of speaking to joking or expressivity. Examples of the former and the latter functions of CS are discussed in Extracts 91–97:

Extract 91

Speakers are looking at photos. D is commenting on the photo of a squirrel that he took.

- (1) D: [...] koło mnie _[RUS] **na sumke** siedzi (.) czeka _[RUS] **tipa**
[...] next to me on bag sits (.) waits sort of
'[the squirrel] sits next to me on the bag (.) [it] sort of waits'
- (2) C: fajnie
cool
'cool'
- (3) D: _[RUS] **dawaj est'** **mne** (.) **czego paseś'**
give to eat me (.) what stare
'give me [something] to eat (.) what are you staring [at]?'
- (4) C: a tutaj co on robił?
and here what he did
'and what did he do here?'
- (5) D: on _[RUS] **tipa** ten (.) jakoś moczy tam żeby kolor swój=
he sort of this (.) somehow wets there so that colour one's own=
'he sort of this (.) somehow makes it wet so that his colour='

The interclausal switch into Russian in line (3) consists of two independent clauses, separated by a brief pause. The change in the base/matrix language from Polish into Russian coincides with a change in the narrative frame, i.e. the speaker differentiates between his own voice and that of a character that he is now portraying by adopting a different language to convey the squirrel's words. There appear to be no identifiable triggers neither in the preceding nor in the following turn that could have contributed to a change in the base/matrix language, hence it is plausible to classify the above example as an instance of discourse-related switching. Once the desired discourse effect is achieved, speakers revert back to using Polish as the base/matrix language.

Another example can be found in Extract 92:

Extract 92

T and S are talking about industrial vehicles.

- (1) T: przezywajo Iveco ^[RUS] **ovečki** ((laughter))
nickname Iveco sheep
'[they] nickname Iveco [as] sheep'

- (2) S: ^[RUS] **Iveko čudo sveta dlja sovremennogo človeka**
Iveco wonder world for modern person
'Iveco - wonder of the world for a modern person'

- (3) T: Vadim gadał że tam ^[RUS] **izoljacija** bardzo drena
Vadim said that there isolation very bad
'Vadim said that isolation is very bad there'

- (4) S: jakie zdjęcia malutkie
what photos small
'what small photos'

The switch into Russian in turn (2) consists of a single main clause and performs a discourse related function. The speaker seems to be quoting an advertising slogan used by the manufacturing company Iveco, or he may have heard it used somewhere else but decided to apply it in this instance. The use of the slogan in Russian, rather than attempting to translate it into Polish not only helps to avoid the unnecessary effort, but also preserve its authenticity and achieve the desired stylistic effect. As turns (3) and (4) demonstrate, the immediately following discourse proceeds with Polish as the base/matrix language.

A related function of CS is exemplified in Extracts 93 and 94, where language alternation coincides with the insertion of a Russian set phrase, i.e. a common expression whose wording is usually not subject to variation. These may include proverbs,

idioms or sayings, which are often colloquial in nature and originate from popular culture:

Extract 93

S is offering something to T.

- (1) S: chcesz?
want
‘[do you] want?’
- (2) T: ni odkażam się
no refuse
‘[I] won’t refuse’
- (3) S: _[RUS] **na haljavu i uksus sladkij**
for free and vinegar sweet
‘even vinegar tastes sweet for free’
- (4) T: to prawda
it true
‘it’s true’
- (5) S: i jeszcze jaka
and more what
‘it is indeed’

The switch into Russian takes place in turn (3) and consists of a single independent clause. It contains a well-known set phrase of a highly colloquial nature, which is further corroborated by the fact that it is listed in Elistratov’s (2002) ‘Slovar’ russkogo argo’ (*Dictionary of Russian Argot*) and Kuzmič’s (2000) ‘Slovar’ narodnoj frazeologii’ (*Dictionary of Folk Phraseology*). While it would be possible to translate the saying into Polish, this would imply losing its stylistic and expressive effect. As turns (4) and (5) demonstrate, after the insertion of the Russian set phrase the base/matrix language goes back to Polish.

A similar example can be found in Extract 94:

Extract 94

- (1) R: ty myśli żeb tobie byłoby dobrze (.) żeb tobie byłoby
you think that you would be good (.) that you would be
wygodny samochód a nie to żeby twoim (0.2)—
comfortable car and not it that your (0.2)—
‘[you] think so that it would be good for you (.) so that the car would be
comfortable for you and not so that—’
- (2) Q: _[RUS] — **babom** —
— women —
‘— women —’
- (3) R: _[RUS] — **babom** podobało się by
— women would like
‘— women would like’
- (4) Q: _[RUS] **potomu što kobel’ kobelja uznaet izdaleka**
because that male dog male dog recognises from afar
‘because one male dog recognises another male dog from afar’
- (5) R: to teraz włosy pomaluj na biały kolor bo dla tej dziewczyny
so now hair dye on white colour because for this girl
bendzi podobać się blondyny
will be like blonds
‘so now dye your hair white because this girl will prefer blonds’
- (6) Q: ja dziewczyny lubia (.) ja nie _[RUS] **gej**
I girls like (.) I not gay
‘I like girls (.) I’m not gay’

The switch into Russian in turn (4) consists of a single independent clause and contains a reference to a well-known Russian song ‘Rybak rybaka uznajot iz daleka’ - *A Fisherman Recognises Another Fisherman from Afar*. The title of the song

has become semi-lexicalised into a set phrase in colloquial Russian with an option of substituting ‘rybak’ by other lexemes, e.g. ‘durak’ - *idiot* or ‘zemljak’ - *fellow countryman*. The speaker might have heard the version with ‘kobel’ used somewhere else, or he may have invented it on the spot, but by employing the Russian phrase he makes a clear reference to the song and appeals to the common cultural background he shares with his interlocutor. The conversation reverts back to the Polish base/matrix language once the desired discourse effect has been achieved.

CS between turns may also be employed when the speaker wishes to foreground a different conversational activity. An example can be found in Extract 95:

Extract 95

C and D are looking at photos.

- (1) D: ja podszedł (.)_[RUS] **koroče vo** tutaj_[RUS] **vo strojka** idzi
 I approached (.) in short here here here construction goes
 ‘I approached (.) basically the construction is under way right here’
- (2) C: _[RUS] **šel na rabotu ustraivat’sja?**
 went to work employ
 ‘did you go to get a job [there]?’
- (3) D: ((laughter)) tam _[RUS] **nelegal’no** kto jej wie czy wlaźł by
 there illegally who her knows if get in would
 ‘who knows if [I] would be able to get in there illegally’
- (4) C: czemu nie (.) trzeba było spróbować
 why not (.) must was to try
 ‘why not (.) [you] should have tried’

The switch into Russian takes place in turn (2), which consists of a single independent clause and appears to perform a stylistic function. When D is showing C a picture of a construction site that he took in Stockholm, C decides to tease him and jokingly asks him whether he went there to look for a job. The humorous nature of this turn is

additionally reinforced by the switch into Russian, which sets it off from the previous turn, where Polish clearly acts as the matrix language. The joke seems to fulfil its desired effect and evokes laughter from D in turn (3), which is then followed by a response. As turn (4) demonstrates, the conversation then reverts back to the Polish base/matrix language.

While in the previous five examples it is possible to attribute a discourse-related or stylistic motivation for switching into Russian, there are also instances where it is problematic to identify an apparent reason for switching. Some examples can be found in Extracts 96 and 97:

Extract 96

- (1) T [...] motor ^[RUS] **naprimer** bendzi już ^[RUS] **odin devjat'** (.) jeszcze
 engine for example will be already one nine (.) more
^[RUS] **dokineš'** na przykład pięć tysięcy to bendzi ^[RUS] **devjanosto šest'**
 will add for example five thousand so will be ninety six
kilovatt
 kilowatts
 '[...] the engine, for example, will already be 1.9 (.) in addition, if you add
 for example five thousand, so it will be 96 kW'
- (2) S: ^[RUS] **ne soblaznjaj** (.) **ja vse ravno Opelja bol'she nikogda ne**
 not tempt (.) I all same Opel more never not
kuplju
 will buy
 'Don't tempt me (.) I will never buy an Opel again'
- (3) T: ja ^[RUS] **vot** ni wiem (.) żeby co tylko nie Golf
 I *emph. particle* not know (.) so that what only not Golf
 'I [personally] don't know (.) anything but Golf'
- (4) S: czemu tak?
 why so
 'why so?'

Although turns (1) and (3) contain several single switches into Russian, their matrix language can clearly be identified as Polish. In contrast, speaker S switches into Russian for the entire duration of turn (2), which consists of two independent clauses. It is difficult to state with certainty why he chose Russian as the base/matrix language for this one turn, as the immediately following discourse proceeds with Polish as the base/matrix language. It is possible to speculate that the speaker is expressing a value judgement and is communicating his negative attitude against this particular car brand, which could have triggered a switch into Russian. However, as with all other examples of CS employed in its expressive function, where language alternation serves to further reiterate the speaker's feelings and sentiments, it is difficult to verify whether the switch is actually driven by expressive motives, particularly when there are no prosodic clues, e.g. an exclamative intonation, or stronger articulation of certain syllables to create additional emphasis. A similar observation applies to the example in Extract 97:

Extract 97

- (1) B: jak ty szybko
 how you quickly
 'how quickly you [came]'
- (2) A: _[RUS] **kak ni stranno v poliklinike očeredej netu**
 how not strange in polyclinic queues is not
 'strange as it may seem, there are no queues at the polyclinic'
- (3) B: tak?
 yes?
 'yes?'
- (4) A: szybko _[RUS] **tak vse** tego (.) wiesz że ja _[RUS] **na četvertom**
 quickly so everything this (.) know that I on fourth
mesjace ciąży
 month pregnancy

‘everything went so quickly (.) Do you know that I am in my fourth month of pregnancy?’

Speaker A switches to Russian in turn (2), which is made up of a single independent clause. Once again, it is difficult to identify the reason behind the switch as the discourse mode appears to be the same throughout the entire exchange. The switched turn does not seem to perform any discourse-related or stylistic functions and the speaker reverts back to using Polish as the base/matrix language with occasional intraclausal switches from Russian in turn (4).

The examples analysed in this section demonstrate that interclausal switching between turns is very similar to interclausal switching within turns. In both types of switching the juxtaposition created by language alternation can be used for discourse-related purposes, such as to differentiate between direct speech and the speaker’s own voice, to set off side remarks and clarifications related to the main narrative, or to make subjective comments or evaluations about what is being said. In addition, switches between turns can be employed to contextualise a different modality, such as shifting from a neutral mode of speaking to joking, as well as to insert longer set phrases, e.g. advertising slogans, idioms or song fragments. However, what is significant is that once the desired discourse or stylistic effect has been achieved, speakers revert back to Polish, which suggests that the departure from the default base/matrix language of the interaction is kept relatively short and the return is fairly predictable, i.e. in the great majority of instances it does not last longer than one turn. Finally, in both types of switching there are instances that must be classified as ambiguous since the reason behind language alternation is not immediately identifiable.

The aim of the next section is to focus on the remaining instances of clauses where Russian can be unambiguously identified as the base/matrix language, but where Polish lexical elements can also be found on the surface level.

6.5 Russian Interclausal Switches with Polish Lexical Elements

The previous two sections dealt with interclausal switches which involve monolingual Russian clauses either within a turn, or between turns. The aim of the present section is to focus on clauses which have Russian as their base/matrix language, but also feature Polish intraclausal switches of the insertional type.

The properties of clauses with Polish as the base/matrix language, which make up the majority of instances in the present data, have been addressed in the two previous chapters, most notably during the analysis of intraclausal switches into Russian in Chapter 5. It has emerged that on the surface level, they consist of both Polish and Russian lexical elements. Likewise, as a result of structural calques from Russian, there are also numerous instances in the data where the underlying abstract structure of the Polish frame is a composite of both languages, which is a strong argument in favour of the Emerging Mixed Code Hypothesis. The aim of this section is to verify whether any of these observations also apply to clauses that have Russian as their base/matrix language.

As far as the surface level is concerned, there are 60 clauses with Russian as the base/matrix language, which contain Polish intraclausal switches of the insertional type, consisting of both single switches and ELIs. This figure can be treated as relatively insignificant given the data presented in the two previous chapters, where Russian lexical items were shown to be inserted into the Polish base with great frequency and regularity. Moreover, the switched items themselves are qualitatively different with regard to several categories. In the overwhelming majority of instances, Russian switches inserted into the Polish base are made up of discourse markers, nouns, verbs, adjectives and ELIs. By contrast, Polish lexical items that are inserted into the Russian base/matrix language consist mainly of adverbs, pronouns, con-

junctions and a few ELIs. Perhaps the most striking difference is that nouns and adjectives, which belong to the lexical category and are inserted into the Polish base with great frequency and success, hardly ever feature in the clauses with Russian as the base/matrix language (only four tokens in total). On the other hand, there are numerous insertions of lexical items belonging to functional categories.

The asymmetry between lexical and functional categories is often brought up in psycholinguistic research, which stems from the fact that the two categories of lexical items display different behaviour patterns. Content words that belong to lexical categories have a more specific and relatively transparent meaning. According to Johanson (2002: 39), the lexicon is the most unstable area, which is most easily influenced and copied in all contact situations. On the other hand, function words, whose meaning is relatively broad and non-specific, are more stable. Moreover, function words are employed more frequently compared to the more specific content words. This ensures that function words are more entrenched in the speakers' lexicons than content words, which further contributes to their relative stability.

Some examples of typical switches that feature in Russian clauses can be found in Extracts 98–101:

Extract 98

A and B are talking about a singer who passed away.

- (1) B: *który zmarł?*
 which died
 ‘the one who died?’
- (2) A: [RUS] *da::* (.) **takoj mužik igral na gitare s usami** (.) **s**
 yes:: (.) such man played on guitar with moustache (.) with
borodoj taki (.) **v očkah**
 beard such (.) in glasses
 ‘yes (.) such a man [who] played the guitar with a moustache (.) such with
 beer (.) with glasses’

- (3) B: taki _[RUS] **tipa** co country śpiewał tam może?
 such kind of what country sang there maybe
 ‘perhaps the one who kind of sang country music?’

The switch in turn (2) is an example of interclausal language alternation since the previous turn and the following turn have Polish as their matrix language. The entire switched turn is expressed in Russian, except for the single insertion from Polish ‘taki’. The employment of the Polish pronoun appears to be completely random, since the speaker could have employed its Russian equivalent ‘takoj’, as she does in the beginning of the turn.

Another example can be found in Extract 99:

Extract 99

- (1) A: on mówi _[RUS] **tut v Jurmalu vo vremja Novoj Volny**
 he says here in Jurmala during new wave
 ‘He says: ‘when you come to Jurmala during the New Wave [festival]’
priedeš’ jak _[RUS] v etom (.) v N’ju Jorke
 come like in this (.) in New York
 [it’s] like in this (.) in New York’

The turn exemplified above consists of two independent clauses and can be classified as an example of interclausal switching for discourse-related purposes. The first one is expressed in Polish, while the second one has Russian as its base/matrix language. The Polish clause is employed to introduce direct speech, whereas the Russian clause is used to convey the actual quotation. It contains one single switch from Polish, namely the adverb ‘jak’.

An example of a Polish ELI can be found in Extract 100:

Extract 100

- (1) B: ja mówia _[RUS] **kak ty do tego _[RUS] s devuškami vstrečalsja?**
 I say how you before that with girls go out

‘I say: how did you go out with girls before that?’

From a pragmatic point of view, the above extract is similar to the previous one. It consists of a clause with Polish as the base/matrix language, which introduces the direct speech, and a clause with Russian as the base/matrix language, which conveys the actual words of the speaker. It can be classified as another example of an interclausal switch for discourse-related purposes, where the contrast created by language alternation is used to separate between the main narrative and the direct speech. Moreover, the Russian clause also contains a Polish ELI ‘do tego’, consisting of a preposition and a pronoun. It is difficult to tell what could have prompted the insertion of a Polish island, as there is no reason to assume that the speaker does not know the equivalent construction in Russian, nor does it perform any discourse, stylistic or grammatical functions.

Finally, an example of a Russian clause with a switched Polish verb can be found in Extract 101:

Extract 101

- (1) A: a teraz pracujesz pracujesz żeby potem ^[RUS] **vyjti na pensiju**
and now work work so that afterwards leave on pension
‘and now [you] work [and] work so that afterwards [you could] retire’
- i pjat’sot šestisot litov** otrzymać
and five-hundred six-hundred litas get
‘and get five or six hundred litas’

The above extract consists of the main clause and two subordinate clauses, though they could also be classified as coordinated infinitives. The main clause is in Polish, whereas the two subordinate clauses are in Russian, except for the conjunction ‘żeby’ followed by the adverb ‘potem’, and the Polish verb ‘otrzymać’, inserted at the very end. As far as the two former lexical items are concerned, they seem to be employed as a continuation from the main clause. As stated earlier, there is a definite preference

for Polish subordinating conjunctions in the present data, hence this finding is not surprising. What is interesting is the use of the Polish verb at the very end of the clause. There is no reason to assume that the speaker does not know the Russian equivalent since it has been employed on several occasions, e.g. A: ‘*Polučili deti diplomy i vse pouhodili*’ – *The children received their diplomas and they all left*. Likewise, there seems to be no discourse-related function that the turn-final switch into Polish could perform, which also seems to apply to all other examples analysed in this section.

The four extracts discussed above demonstrate that Russian clauses, which may or may not perform an identifiable discourse-related function, are kept monolingual in the great majority of instances, i.e. 217 tokens. However, there are further 60 tokens of clauses with Russian as the base/matrix language, which also contain intrasentential switches into Polish in the form of single lexical items and ELIs. As demonstrated in Extracts 98–101, the tendency is to insert either verbs, adverbs and ELIs, or lexical items that belong to functional categories, most often conjunctions. These examples demonstrate that clauses with Russian as the base/matrix language may also contain both Russian and Polish lexical elements on the surface level, though in quantitative terms Polish intraclausal switches are not as significant as Russian switches analysed in the previous chapter. Nonetheless, their presence indicates that speakers never fully depart from the bilingual mode of speaking for their in-group interactions and readily go back to mixing even when the matrix language changes from Polish, which seems to be the default choice, into Russian. This finding can be taken as another piece of evidence in favour of the Emerging Mixed Code Hypothesis.

While intraclausal switches on the surface level are characteristic not only of Polish, but also Russian clauses, there is a significant difference between the two when it comes to the abstract level. Recall from the previous chapter that there are clauses which appear to originate from Polish on the surface level, but due to the presence

of structural calques, their abstract structure is in fact a composite of Polish and Russian languages. The analysis of both monolingual and bilingual Russian clauses shows that there is no evidence of Polish influence on the abstract level. This finding is significant since it demonstrates that we are dealing with contact-induced structural change, which can be regarded as unidirectional, i.e. from Russian (L2) to Polish (L1). From the sociolinguistic perspective this further corroborates the favourable status attributed to Russian and the prestige that it enjoys among the respondents. From the psycholinguistic perspective this suggests that Russian structures are well entrenched in the lexicons of the speakers and are sometimes preferred over Polish ways of lexicalising given concepts, which is conditioned by the sociolinguistic reality of its use. Since there is a direct link between the frequency with which a speaker encounters a given structure and its level of cognitive entrenchment, the fact that Russian is more often used in a monolingual mode, while Polish is nearly always used in a bilingual mode ensures that Russian structures remain well entrenched and possibly more easily activated in language production, which reinforces its influence on the abstract level in the Polish utterances.

Following the discussion presented above it can therefore be concluded that in the present data there is evidence of convergence on the abstract level, but only of the regional variety of Polish towards Russian, and not in the opposite direction. However, this finding should be taken with a degree of caution. Since there are relatively few examples of clauses with Russian as the base/matrix language, more data would be required in order to arrive at a substantial conclusion.

6.6 Conclusions

To summarise, the aim of the present chapter was to test the evidence in favour of the second hypothesis, namely the ‘classic’ CS hypothesis, which states that there

is a clear preference for one language of interaction, that there is a change in the base language from the switching point onwards where the return to the previous language of interaction is not predictable, that it is possible to assign at least a tentative discourse-related function that CS is said to perform and that there is little or no evidence of convergence. The analysis focused on interclausal switches, since they tend to be regarded as prototypical examples of ‘classic’ CS.

One of the main findings of the present chapter is the very fact that interclausal switches do occur alongside intraclausal switches, which suggests that the two types of language alternation, namely CS and CM co-exist. However, it should be noted that, with a total of 322 instances, interclausal switches are not as numerous as intraclausal switches, which amount to 3,521 tokens. Some asymmetry in the number of tokens is to be expected, since intraclausal switches tend to consist of a single lexical item and can therefore be employed much more often. Nevertheless, the number of tokens indicates a strong preference towards switches of the intraclausal type and a relatively low number of switches of the interclausal type.

In order to test the ‘classic’ CS hypothesis the analysis focused on both interclausal switches within a turn and interclausal switches between turns. Only one of the premises of the ‘classic’ CS hypothesis is, to some extent, borne out by the data analysis, namely that language alternation can perform an identifiable discourse function. However, as noted earlier, any potential functions that CS is said to perform are based on the subjective interpretation of the researcher, hence they should be treated as plausible guesses rather than empirically based. The most frequently attested functions tend to be linked to creating turn-internal structure and include differentiating between direct speech and the speaker’s own voice, setting off side remarks and departures from the narrative frame, making comments and evaluations about what is being said, signalling a change of topic or inserting an idiom, advertising slogan or another set phrase. These findings are consistent with the descriptions of

‘classic’ or prototypical CS given in the existing literature.

However, the data analysis also shows that in roughly half of all instances of interclausal switching even a tentative discourse-related function cannot be identified and they are therefore classified as ambiguous. One possible explanation for this finding is the fact that back and forth language alternation between Polish and Russian is the accepted mode of speaking in the speech community studied in this thesis, which means that speakers may occasionally exploit the juxtaposition between the two languages to structure their utterances and to contextualise certain verbal activities, but, as with intraclausal switching, they do not necessarily switch with a particular discourse or stylistic effect in mind.

The remaining three premises of the ‘classic’ CS hypothesis are not supported by the data. Contrary to what the first premise stipulates, the bilingual mode of speaking is the preferred type of language interaction. While Polish tends to act as the base/matrix language, frequent insertions from Russian are regarded as the norm and amply feature in the discourse of all the speakers in the data. As far as the second premise is concerned, it has been demonstrated in the present chapter that speakers may momentarily depart from Polish as the base/matrix language to Russian in order to contextualise a discourse-related activity or achieve a desired stylistic effect; however, this is usually achieved within a single turn. This suggests that the return to the default language of interaction can be regarded as predictable, which is the opposite of what the ‘classic’ CS hypothesis stipulates.

Finally, the final premise of the ‘classic’ CS hypothesis has already been proven false in the previous chapter, where structural convergence in the form of calques has been addressed in greater detail. Following the analysis presented in the present chapter, it can be further concluded that there is no evidence that would suggest the influence from Polish on the underlying abstract structure in clauses with Russian as the base/matrix language. There are instances where Polish lexical items are inserted

on the surface level, though these insertions are to some extent qualitatively different from the Russian insertions into the Polish base. Most notably, Russian intraclausal switches mainly include lexical categories, as well as ELIs, whereas Polish intraclausal switches may also include lexical items that belong to functional categories, most notably conjunctions. While in quantitative terms Polish intraclausal switches are not as significant as Russian switches, their presence indicates that speakers never fully depart from the bilingual mode of speaking for their in-group interactions and readily go back to mixing.

It can be concluded that the findings of this chapter offer little support for the ‘classic’ CS hypothesis. While interclausal switching of the alternational type does feature in the present corpus, its presence is much less regular and significant compared to that of the intraclausal switching of the insertional type. On the other hand, the analysis of grammatical features of Russian clauses provides some further support for the Emerging Mixed Code Hypothesis, since it further confirms that speakers never fully depart from the bilingual mode of speaking.

Chapter 7

Discussion and Conclusions

Chapters 4, 5 and 6 present an overview of three main types of language alternation that feature in the present data, namely switching at discourse markers, intraclausal switching of the insertional type, which involves single switches and ELIs, and interclausal switching of the alternational type. Each of them reveals a different aspect of the nature of language alternation prevalent in the Polish community in Lithuania. The aim of the present chapter is to discuss the significance of the findings in the context of the present thesis and to decide to what extent these different types of language alternation can be reconciled with each other.

7.1 Prototypical Code-Switching or Code Mixing?

The main aim of the present study was to evaluate whether the pattern of language alternation prevalent in the Polish community in Lithuania can be classified as an instance of an emerging mixed code, or whether it displays more features characteristic of the ‘classic’ or prototypical CS. The possible outcomes of language contact and the principal differences between different models of language alternation phenomena were discussed in greater detail in Chapter 2, hence they are only briefly recapitulated here

Classic Code-Switching	Code-Mixing/ Emerging Mixed Code	Fused Lect/Fusion
<ul style="list-style-type: none"> • preference for one language of interaction at a time • interclausal switches of the alternational type for discourse or stylistic purposes • locally meaningful insertions of intraclausal switches 	<ul style="list-style-type: none"> • significant increase in intraclausal switches of the insertional type (both single switches and larger constituents) • presence of convergence on the grammatical and phonological level • loss of salience and local meaning of switches due to increased frequency and regularity • code-mixing regarded as the default and preferred mode of interaction • it is not always possible to identify the base/matrix language • no preference for one language of interaction at a time 	<p>Same as mixing plus:</p> <ul style="list-style-type: none"> • functional specialisation (grammaticalisation) of certain constituents from the embedded language, such as discourse markers and conjunctions, as a result of frequent and regular usage in the base language

Table 7.1: Overview of possible outcomes of language contact

in Table 7.1.

As stated in the introduction, the presence of one form of language alternation does not exclude the other, for several types of language alternation may co-exist within a given speech community. It is generally agreed that there is a tendency to move from CS to CM and eventually a fused lect, but not in the opposite direction. Moreover, during the transitional process, there are phases in which the ‘older’ patterns of CS, namely locally meaningful switches of the alternational type, co-exist with the ‘newer’ patterns of CM, namely frequent switches of the insertional type. As Table 7.1 demonstrates, the transition to CM is typically spearheaded by an increase in intraclausal switching of the insertional type, which is restricted not only to single lexemes, such as nouns and verbs, but spreads to larger constituents, such as noun, verb and prepositional phrases, as well as idioms and collocations (Auer 1999: 320–321). Another important feature is the presence of convergence, which can manifest itself in the underlying grammatical level via single-word and structural calques and contribute to the development of the composite matrix language.

The final step on the continuum is the transition from CM to a fused lect, which presupposes functional specialisation of certain constituents from the embedded language, such as discourse markers, as a result of frequent and regular usage in the base language. As Table 7.1 demonstrates, apart from grammaticalisation, the characteristics for fused lects and CM are pretty much the same, which suggests that the usefulness of this category is questionable. Fused lects are said to be based on the presence of only one additional diagnostic criterion, which raises the question of whether it really justifies the introduction of an entire new category. Although in the present study grammaticalisation affects only DMs, in theory this process could spread to other categories as well. For example, if there was evidence that aspectual usage of Polish motion verbs always follows the rules of Russian morpho-syntax, then this could be classified as another instance of grammaticalisation. Something like

this would not be captured by CM at all, which in turn justifies maintaining the distinction between CM and fused lects, with the latter label reserved to account for the more conventionalised structural regularities arising as a result of language mixing.

The analysis presented in the previous chapters demonstrates that the bilingual mode of speaking observed in the Polish community in Lithuania is made up of two different ways of juxtaposing the regional variety of Polish and Russian, with some occasional single switches from Lithuanian. On the one hand, there are numerous instances of intraclausal switches of the insertional type, which do not seem to carry any local meaning and which involve not only single switches, but also include larger constituents. Following Auer's (1999) theory, this type of switching would be classified as an instance of the 'newer' patterns of CM. On the other hand, there are cases of interclausal switches of the alternational type, which often perform a discourse-related function, where CS is employed to create internal structure within or between turns and which would be classified as the 'older' patterns of CS. The relevant issue to be considered at this stage would be to what extent does one type of switching prevail over the other?

From a purely quantitative perspective, intraclausal switches of the insertional type clearly outnumber interclausal switches of the alternational type and after discourse markers, constitute the preferred type of language alternation. In the present corpus there are 3,521 instances of the former, and only 322 instances of the latter. On the basis of this quantitative information alone it could be argued that the data display compelling evidence in favour of CM. However, quantitative prevalence alone is not enough to support the Emerging Mixed Code Hypothesis and other factors should be taken into consideration.

In addition to quantitative evidence, it is also important to take into account qualitative findings. The data analysed in Chapter 5 provide ample examples of intraclausal

switches of the insertional type. Single switches from Russian are not only extremely common, but they also feature in the discourse of each and every speaker in the present corpus. In addition to being frequent and wide-spread, they are also non-recurrent and do not carry any apparent local meaning. Likewise, there is evidence which suggests that switching has further spread to include larger constituents, which are referred to as ELIs. While not as numerous as single switches, ELIs can also be described as non-recurrent, they feature in the discourse of each and every speaker in the present corpus and thus they can be regarded as relatively wide-spread. The characteristics of intraclausal switches outlined here show that frequent Polish-Russian language mixing is accepted in the Polish speech community analysed in this thesis and that it can be regarded as the default mode of speaking.

While frequency of occurrence, acceptance within the Polish speech community and apparent lack of local meanings constitute one piece of evidence for the Emerging Mixed Code Hypothesis, further evidence comes from phonetic and morphological characteristics of switched items inserted into the Polish base. As demonstrated in Chapter 5, intraclausal switches tend to retain their Russian pronunciation and are inserted with their original late system morphemes. To a great extent this is due to the fact that the two languages in contact are typologically closely related and congruent to a high degree, which facilitates insertion of Russian lexical items and thus contributes to language mixing. However, the fact that late system morphemes from both languages are permitted can be taken as another piece of evidence for the language mixing hypothesis, for it shows that simultaneous employment of both sets of morphemes is accepted as grammatically correct by the speakers.

Another piece of evidence that points towards language mixing is the presence of convergence on two different levels of the regional variety of Polish towards Russian, which was described in greater detail in Chapter 5. The most prominent examples of convergence are calques, which may consist of a single lexical item or a larger

constituent. Their defining characteristic is that the underlying abstract structure originates from the embedded language, namely Russian, but the surface level is that of the base/matrix language, namely Polish. The presence of calques is very important for the Emerging Mixed Code Hypothesis, since it proves that Polish and Russian are combined not only on the surface level, but that at least part of the abstract structure of the frame itself is a composite of the two languages.

At this stage it is difficult to tell how advanced this convergence is. As demonstrated in Chapter 5, there are 155 instances of single word calques and 83 instances of structural calques in the present corpus. Compared to other phenomena of language contact, e.g. intraclausal switching, these figures can be regarded as fairly insignificant. However, there are two important aspects that should be borne in mind. The first one is that calques are notoriously difficult to identify, which means that there may be more instances of calquing which have not been spotted during the initial analysis. The second aspect is that the present corpus is relatively small, thus more data would be required in order to make a proper assessment. Nonetheless, the very fact that the presence of Russian is not limited to insertions on the surface level constitutes a strong piece of evidence in favour of the mixed code hypothesis.

Additional evidence comes from the fact that there are no metalinguistic comments on the fact that the speakers are code-mixing. While some informants did include brief metalinguistic digressions, they mainly commented on the topics discussed in their conversations or apologised for the swearwords. However, not a single speaker made any remarks about the purely linguistic aspects of their conversations. This suggests that frequent alternation between Polish and Russian is regarded as the unmarked choice among members of the speech community, studied in this thesis, and can be taken as another piece of evidence supporting the Emerging Mixed Code Hypothesis.

This piece of evidence can be further supplemented with data from the analysis of

interclausal switches of the alternational type. As demonstrated in Chapter 6, there are 322 instances of this type of language alternation in the present corpus. However, in roughly half of the instances, language alternation does not seem to perform any apparent discourse-related or stylistic functions. While it is possible to speculate that in some of these instances switching is associated with expressive functions of CS, these assumptions are plausible guesses, rather than empirically based. One possible explanation as to why there are instances of switching between clauses which carry no apparent local meanings is that in the present data, the bilingual mode of interaction is regarded as the norm, which means that speakers no longer switch with a particular discourse or stylistic effect in mind. This would also suggest that not only intraclausal switching, but also interclausal switching is becoming less and less functional.

Finally, some examples of clauses with Russian as the base/matrix language demonstrate that the process of intraclausal switching also takes place in the opposite direction, i.e. Russian base – Polish switch. While in quantitative terms Polish intraclausal switches are not as significant as Russian switches, their presence further confirms the prevalence of the bilingual mode of speaking within the Polish speech community, analysed in the present data. It shows that speakers readily resort to CM even in those utterances that are intended to be monolingual and that take place in a different base/matrix language. However, it should be noted that there are relatively few examples of the latter, hence the extent of this observation cannot be properly verified.

To summarise the argument so far, strong evidence in favour of the Emerging Mixed Code Hypothesis comes from the quantitative prevalence of intraclausal switches, as well as from the fact that they are non-recurrent, regularly feature in the discourse of all of the speakers in the data, and can appear with both Polish and Russian morphological endings. Moreover, additional support is provided by the examples of

convergence of the regional variety of Polish towards Russian on the grammatical level, where the underlying abstract structure of the matrix frame can be regarded as the composite of both languages. Finally, the fact that there are instances of interclausal switches where language alternation does not carry any apparent discourse-related or stylistic functions, as well as the fact that intraclausal switches from Polish feature in clauses with Russian as the base/matrix language can also be regarded as additional pieces of evidence, though they should be treated with a degree of caution and need further verification.

Evidence in favour of the ‘classic’ CS hypothesis is presented in Chapter 6, where interclausal switches of the alternational type are investigated in greater detail. Overall, there are 322 instances of this type of language alternation in the present corpus, which means that the relevance of ‘classic’ or prototypical CS cannot be immediately dismissed. While in half of the instances it is difficult to attribute an apparent discourse-related or stylistic function that switched clauses could perform, it can certainly be done with regard to all the remaining examples, which supports the prevalent assumption that in ‘classic’ CS language alternation is locally meaningful. It has been demonstrated in Chapter 6 that switching between clauses may be employed to differentiate between direct speech and the speaker’s own voice, set off clarifications and side remarks, or make subjective comments and evaluations about what is being said.

Additional evidence in favour of the ‘classic’ CS hypothesis comes from the analysis of the directionality of switching. In the examples of interclausal switches that occur within a turn, in the overwhelming majority of instances a switch proceeds from Polish to Russian. As demonstrated in the existing literature, the direction of the switch can be indicative of the speaker’s dominant language, since language alternation tends to start with the dominant language and move towards the subordinate language (Edwards and Dewaele 2007: 230). Drawing on this observation would suggest that

Polish acts as the dominant, and hence the preferred language of interaction in the present data, whereas Russian can be regarded as the subordinate and therefore secondary language. While this does not mean that speakers orient themselves towards a monolingual mode of interaction, this finding partly suggests that one of the participating codes is regarded as the default option.

This observation is further supported by the overwhelming preference for Polish as the base/matrix language of the interaction, which is further corroborated by the fact that in the present corpus the overall number of Polish orthographic words still prevails over Russian by a large margin. While there is some counter evidence, namely those instances when the base/matrix language appears to be ambiguous, as well as examples where Russian acts as the matrix language while Polish takes the role of the embedded language, it is not numerous enough to dismiss the overall prevalence of Polish in the present data, as well as a clear preference for the Polish base.

On the basis of the above discussion it can be concluded that a possible move from CS to a mixed code has taken place in the Polish speech community investigated in this thesis. Judging from the evidence, CM can be regarded as the dominant type of language alternation attested in the present data. However, switching patterns characteristic of 'classic' CS can also be identified, which suggests that transition into an advanced mixed code is not fully complete yet. This is further supported by the fact that Polish is the default base/matrix language chosen by the speakers. At this stage it is difficult to predict whether the pattern of language alternation documented in this study is likely to develop further, or whether it will remain fossilised in its current state for a prolonged period of time.

7.2 Typological Proximity and its Implications

One of the issues that needs to be addressed in greater detail is the typological

proximity between the two languages involved in CM and its implications for the data analysis presented in this thesis, as well as theories on language mixing. Both Polish and Russian belong to the Slavic family of languages, and although Polish is a West Slavic language, while Russian is an East Slavic language, they nevertheless exhibit strong lexical and grammatical similarities. On the lexical level, they share numerous cognates, namely many functional parts of speech, such as prepositions, conjunctions and adverbs, as well as certain verb forms, nouns and pronouns (cf. ‘z’ - *from*, ‘na’ - *on*, ‘i’ - *and*, ‘tam’ - *there*, ‘byli’ - [*they*] *were*, ‘dom’ - *house*, ‘on’ - *he*, ‘my’ - *we*). On the morpho-syntactic level, they both follow a basic SVO word order and display a fairly elaborate system of declensions with some overlapping inflectional and derivational morphemes. The Polish-Russian data analysed in this thesis confirm that the lexical overlap between the two languages facilitates frequent and regular insertional switching on the lexical level. In addition, the presence of larger syntactic calques and ambiguous matrix language constructions suggests that structural similarity between the two contact languages can be regarded as a precondition for greater structural influence as well.

Although typological characteristics outlined above facilitate mixing on both lexical and structural levels, as demonstrated in Chapter 5, the proximity between the two languages poses a problem for the MLF model. While the System Morpheme Principle predicts that in mixed constituents the so-called ‘late’ system morphemes must come from the matrix language, the data analysis reveals that when it comes to Polish-Russian language alternation, Russian late system morphemes are not only permitted, but they are actually preferred. This scenario is a complete opposite of what the MLF model stipulates. Similar issues with applying Myers-Scotton’s model to closely related languages or dialects have already been reported in some of the existing literature, e.g. Francescini (1998) for CS between the Lombard dialect and Standard Italian, Berruto (2005) for Italian and Italo-Romance dialects. The question that arises from this discussion is why the MLF does not hold for typologically closely

related languages, such as Polish and Russian.

The MLF model is based on the idea of asymmetry between the languages involved in CS: the dominant matrix language is set to supply the main grammatical frame, including outsider late system morphemes, while the embedded language contributes some of the lexical content. The problem with this set-up is that it holds best when the languages are typologically distant. In fact, the more typologically distant the languages are, the more salient this asymmetry is likely to be, and the closer the languages are, the less salient the asymmetry is. Consequently, structural overlap is eased when a grammatical category from one language is fairly analogous to the corresponding category in the other language, which is often the case with Polish and Russian ‘late’ system morphemes.

We know from the psycholinguistic research that in bilinguals with typologically similar languages representations can be shared at all levels where there is storage: the conceptual level, the syntactical level, the discourse level, the gesture level, the syllable level and the phoneme level. We also know that occasional CS is fundamentally different from habitual language mixing, where the two languages are both highly activated (de Boot et al. 2009: 100). What follows is that if the structures of the two highly activated languages involved in habitual mixing correspond typologically, then this similarity will facilitate a direct mapping of morphemes from one language to another, including what Myers-Scotton refers to as ‘late’ system morphemes. It can therefore be concluded that typological distance inhibits the transfer of inflectional morphology, whereas typological similarity facilitates it, which puts Myers-Scotton’s model in jeopardy.

The observations presented above suggest that typological proximity/distance play a crucial role when it comes to theories on language mixing. The problem with the MLF model in particular is that its main premise is that the same principles underlie all language contact phenomena and that the proposed model is supposed to be

universal. While in one of the earliest formulations of the model Myers-Scotton states explicitly that it does not apply to dialect contact situations (1997: 17), this exclusion should not affect the Polish-Russian data. The two languages are typologically closely related, but far from dialects of the same language.

It is suggested that in studies of CS and other type of relevant phenomena that arise as a result of language contact, such as material and structural borrowing, typological differences and similarities between languages should be given more importance than they often receive. Bearing in mind that mixing patterns and other phenomena of contact-induced language change are conditioned by the typological characteristics of the language pairs involved, it would be more reasonable to compare mixing outcomes involving either same languages, or typologically similar languages at the very least, rather than aim for developing a universal set of rules or syntactic constraints on CS. Since many syntactic characteristics are language-specific, different language pairs may be subject to different rules of morpho-syntactic integration, which will ultimately be conditioned by the degree of typological closeness or distance between the two languages involved.

7.3 From Code Mixing to a Fused Lect

According to the existing literature, discourse markers occupy a special status when it comes to the transition from CS to CM, and especially from CM to a fused lect, which is an even further development from mixing. The difference between CM and fused lects is mainly a grammatical one, which becomes visible at a deeper structural level only. While language mixing by definition allows variation (languages may be juxtaposed, but they need not be), the use of one code or the other for certain constituents is obligatory in fused lects, for it becomes part of their grammar, and speakers no longer have a choice since variation is lost (Auer 1999: 321). Lexical items that are affected at the earliest stages of the transition process are relatively

unbound elements of grammar, such as discourse markers, conjunctions, and certain adverbials, which typically act as utterance modifiers (Matras 1998; Auer 1999). Unbound elements can be borrowed with greater ease than bound elements for several reasons. Bound morphemes are not only structurally integrated and appear as part of other lexical items, but they also express a more general meaning compared to such unbound elements as nouns or adjectives. Moreover, they are said to be integral parts of the positional frames with which speakers bind content words; consequently, they do not have to compete for selection the way that content words do. This functional property contributes to their regular and frequent use, which further reinforces their relative stability in the speakers' mental lexicon.

One of the reasons why discourse markers and other utterance modifiers are affected at the earliest stages of transition stems from the fact that much like bound morphemes, they are highly recurrent and occur rather frequently in spoken discourse. Frequency of occurrence reinforces their level of entrenchment in the mental lexicon, which means that their selection in the process of language production becomes relatively automatic. However, unlike bound morphemes, discourse markers are unbound and can be employed as stand-alone items. Moreover, they are often found at clause peripheral positions, which reinforces their salience and discourse functions that they perform. All of these characteristics make them more prominent in language contact situations and contribute to their high borrowability ranking. As far as the present data are concerned, discourse markers constitute one group of lexical items that could have potentially been subject to the transition process described above.

As discussed in Chapter 4, Russian unbound elements in the form of discourse markers regularly feature in the discourse of all the speakers. With a total number of 5,268 tokens, they constitute the preferred type of switching in the present data. The most frequently employed discourse markers are 'koroče' – lit. 'shorter' – *in short*, 'tipa' – *sort of*, 'prosto' – *simply*, as well as the discourse particles 'nu' and 'vot'. As

demonstrated in Chapter 4, the latter two lexical items are particularly frequent and numerous. Not only do they feature in the discourse of each and every speaker in the data, they are employed without any significant competition from Polish (apart from seven instances of ‘no’ in the discourse of two speakers). Moreover, the insertion of these discourse particles is non-consequential on the level of language choice, which further supports the view that they are no longer regarded as ‘foreign’, but rather as integral, and hence obligatory elements. These observations provide solid evidence in favour of the Mixed Code/Fusion Hypothesis and allow to suggest that the two discourse particles have been grammaticalised, i.e. ‘fused’ as part of the emerging Polish-Russian mixed code.

The same observation can be applied to ‘koroče’ – lit. ‘shorter’ – *in short*, ‘tipa’ – *sort of* and ‘prosto’ – *simply*. As demonstrated during the data analysis in Chapter 4, these three lexical items are employed without any immediate equivalents from Polish, which suggests that they may be the only and therefore the obligatory constituents. Although they are not as numerous as the two discourse particles, on the basis of the available data it can be concluded that the three discourse markers have been subject to the process of ‘fusion’ and undergone the transition into a fused lect on a par with ‘nu’ and ‘vot’.

Finally, the empirical evidence presented in Chapter 4 suggests that the hypothesis of ‘fusion’ can be further extended to include two Russian swearwords, ‘blin’ and ‘bljad’’. Much like discourse markers and particles, the two swearwords contribute to the organisation of the interaction and the structuring of verbal exchange. Moreover, they appear in the present corpus with great frequency and regularity in the discourse of the majority of speakers without any immediate competing forms or any syntactic restrictions and can be employed as stand-alone items in the Polish base. These findings suggest that they are no longer regarded as ‘foreign’ elements, but rather as an integral part of an emerging Polish-Russian mixed code, or in other words, they

have been ‘fused’ alongside Russian discourse markers and particles.

Judging from the findings presented above, it appears that some lexical items, namely discourse markers, discourse particles and swearwords have been subject to an ever further development from CM and transitioned into a fused lect. Following all the findings presented in this section, an emerging Polish-Russian mixed code, attested in the present data, can be described as based on the Polish underlying morpho-syntactic frame with surface level originating from both Polish and Russian, and the majority of discourse-structuring functions carried out exclusively by Russian lexical elements.

7.4 Drawbacks of the Present Study and Directions for Future Research

Some of the limitations of the analysis presented in this thesis have already been highlighted in the preceding chapters. The aim of this section is to further expand on some of these issues and outline possible areas for future research.

One of the first limitations that should be mentioned here is related to the size and organisation of the present corpus. Compared to other existing corpora, the corpus of Polish-Russian data, which serves as the basis for the present study is considered relatively small. This means that some of the linguistic phenomena could not be investigated in as much quantitative detail as one would hope. They include the presence of material and semantic borrowing, as well as structural convergence, which would have been more accurate and substantial if a greater number of tokens could be collected. Another drawback of the present corpus is the fact that it is untagged as far as lexical categories are concerned, which makes it problematic to offer adequate quantitative information in some instances, e.g. provide the proportion of Polish

nouns in order to evaluate the extent of the switched Russian nouns. Following these observations, one of the possible areas for future research would be the expansion of the present corpus with additional data, as well as the creation of a system of relevant grammatical tagging, which would facilitate quantification, analysis and retrieval of the required lexical items.

The second limitation relates directly to the data analysis presented in this thesis. The discussion of results, which features in Chapters 4, 5 and 6, is primarily concerned with investigating Russian switches in the Polish base. Consequently, very little attention is paid to instances where Polish switches are inserted into the Russian base, or cases where the base/matrix language is ambiguous. In order to provide a more comprehensive analysis of the Polish-Russian language alternation, and to assess the prevalence of CM, it would be important to focus on these two areas in greater detail.

As far as directions for future research are concerned, it is important to single out two pivotal areas. The first one is a comprehensive phonetic analysis of the regional variety of Polish spoken in Lithuania. Not only is there very little information available about the phonetic features of the Polish language spoken in Lithuania, the existing research tends to be outdated and relatively superficial, thus it would benefit from a more up-to-date investigation. Naturally the most crucial question related to contact phonetics would be whether there is any evidence of convergence between the local Polish and Russian phonetic systems? It has been claimed in some of the existing literature that in language mixing some parts of the phonetic system are less likely to be adapted than others (cf. van Gijn 2009). If so, does this observation hold for the Polish-Russian contact situation and what are the levels that have been affected?

The second area for future research is related to psycholinguistic aspects of language alternation. The corpus-based analysis presented in this thesis raises a number of questions which cannot be answered adequately without relying on experimental

data. One of these questions is whether instances of structural borrowing, i.e. calques, are already stored in the speakers' mental lexicon or whether they are produced online. The fact that they occur relatively frequently throughout the corpus suggests that they are already stored. However, in order to confirm this further participants could be instructed to perform a lexical decision task based on deciding whether the target lexical item is a 'real' Polish word or not. Likewise, the corpus-based analysis reveals that there may be some instances of priming in the present data. This could be verified further by asking the participants to take part in a fragment completion or a picture description task to investigate whether the syntactic properties of the prime sentence will influence participants' syntactic choices.

7.5 Conclusions

This thesis started with a broad overview of the historical and linguistic background of the Polish community in Lithuania, followed by a summary of the existing literature on CS, CM and 'fusion'/fused lects. The process of data collection and transcription, as well as the methodology for recruiting informants for the present study were described in detail in Chapter 3. The data analysis, presented in Chapters 4, 5 and 6 focused on describing the types of language alternation that feature in the present corpus, with an overall aim of evaluating which types of language alternation are attested in the data and whether the pattern of Polish-Russian language alternation can be qualified as an instance of an emerging Polish-Russian mixed code.

A summary of the main findings of this thesis would be that, in fact, all three types of language alternation feature in the present corpus. However, the major finding has to be the evidence supporting the code mixing hypothesis, which immediately manifests itself in the significant presence of Russian intraclausal switches, made up of single lexical items and embedded language islands inserted into the Polish base. As a

result of their wide spread and frequency, insertional switches of this type have lost their immediate local meanings and the pattern of frequent Polish-Russian language alternation, referred to as the Polish-Russian mixed code, has become the accepted mode of speaking within the Polish speech community.

Another important finding would be that the contact between the two languages goes beyond the surface level, for in some instances it has also affected the underlying grammatical structure, which can be described as a composite of Polish and Russian. This process is exemplified with the presence of calques, which may affect single lexical items, most notably verbs, as well as larger structures. It has also been pointed out that the process of structural convergence is unilateral, i.e. it is the regional variety of Polish that is converging towards Russian, and not vice versa.

In addition to CM, the data also display limited evidence for an emergence of a fused lect, where the use of one of the languages for certain grammatical constituents has become obligatory. With regard to the present data, lexical items which appear to display characteristic features of 'fusion'/fused lect are the Russian discourse particles 'nu' and 'vot', the Russian discourse markers 'koroče' – lit. 'shorter' – *in short*, 'tipa' – *sort of* and 'prosto' – *simply*, as well as the Russian swearwords 'blin' – *damn* and 'bljad'' – *whore*. All of them appear throughout the corpus with great frequency and regularity, and have no immediate competition from Polish, which suggests that they have become grammaticalised as part of the Polish-Russian mixed code.

Finally, the presence of 'older' patterns of language alternation, namely 'classic' CS, has also been attested in the present data. During the data analysis it was demonstrated that in some instances speakers choose to exploit the juxtaposition created by the contrast between Polish and Russian clauses in order to contextualise a discourse-related activity, such as introducing a shift in the narrative frame or including a subjective comment. The presence of 'classic' CS does not call into question the Emerging Mixed Code Hypothesis, since it merely confirms that the three types of

language alternation co-exist within the Polish-Russian data, though at this stage it is CM that appears to prevail.

Bibliography

- Aijmer, K. (2002). *English Discourse Particles: evidence from a corpus*. John Benjamins Publishing.
- Alfonzetti, G. (1998). The conversational dimension in code-switching between Italian and dialect in Sicily. In Auer, P., editor, *Code-Switching in Conversation: language, interaction and identity*, pages 180–210. London: Routledge.
- Andersson, L. and Trudgill, P. (2007). Swearing. In Monaghan, L., editor, *A Cultural Approach to Interpersonal Communication*, pages 195–199. Oxford: Blackwell.
- Angermayer, P. (2005). ‘*Speak English or what?*’ - *Codeswitching and interpreter use in New York Small Claims Courts*. PhD thesis, New York University, Department of Linguistics.
- Atkinson, J. and Heritage, J. (1984). *Structures of Social Action (Studies in Conversation Analysis)*. Cambridge: Cambridge University Press.
- Auckle, T. and Barnes, L. (2011). Code-switching, language mixing and fused lects: emerging trends in multilingual mauritius. *Language Matters*, 42(1):104–125.
- Auer, P. (1984). *Bilingual Conversation*. John Benjamins Publishing.
- Auer, P. (1995). The pragmatics of code-switching: a sequential approach. In Milroy, L. and Muysken, P., editors, *One Speaker, Two Languages: cross-disciplinary*

perspectives on code-switching, pages 115–135. Cambridge: Cambridge University Press.

Auer, P. (1998). Introduction: Bilingual conversation revisited. In Auer, P., editor, *Code-Switching in Conversation: language, interaction and identity*, pages 1–24. London: Routledge.

Auer, P. (1999). From codeswitching via language mixing to fused lects. Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism*, 3(4):309–332.

Auer, P. (2001). Code-switching: discourse models. In Mesthrie, R., editor, *Concise Encyclopedia of Sociolinguistics*, pages 443–446. Oxford: Elsevier.

Auer, P. (2007). The monolingual bias in bilingualism research, or why bilingual talk is (still) a challenge for linguistics. In Heller, M., editor, *Bilingualism: a social approach*, pages 319–339. New York: Palgrave Macmillan.

Auer, P. and Eastman, C. M. (2010). Code-switching. In Jaspers, J., V. J. and Östman, J.-O., editors, *Society and Language Use*, pages 84–112. John Benjamins Publishing.

Auer, P. and Muhamedova, R. (2005). ‘Embedded language’ and ‘matrix language’ in insertional language mixing: some problematic cases. *Rivista di Linguistica*, 17(1):35–54.

Avanesov, R. (1956). *Фонетика современного русского литературного языка*. Издательство Московского Университета.

Avina, N. (2006). *Родной язык в иноязычном окружении : на материале русского языка в Литве : монография*. Moscow, Vilnius: Elpis.

Backus, A. (1992). *Patterns of Language Mixing: a study in Turkish-Dutch bilingualism*. Wiesbaden: Otto Harrassowitz.

- Backus, A. (2001). The role of semantic specificity in insertional codeswitching: evidence from Dutch-Turkish. In Jacobson, R., editor, *Codeswitching Worldwide II*, pages 125–154. Berlin; New York: Mouton de Gruyter.
- Backus, A. (2003). Units in codeswitching: evidence for multimorphemic elements in the lexicon. *Linguistics*, 41(1):83–132.
- Berruto, G. (2005). Dialect/standard convergence, mixing, and models of language contact: the case of Italy. In Auer, P., Hinskens, F., Kerswill, P., editors, *Dialect change. Convergence and divergence in European languages*, pages 81–95. Cambridge University Press.
- Bhatia, T. and Ritchie, W. (1996). Bilingual language mixing, universal grammar, and second language acquisition. *Handbook of Second Language Acquisition*, pages 627–688.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive psychology*, 18(3):355–387.
- Bolinger, D. (1975). *Aspects of Language. 2nd Edition*. New York: Harcourt Brace Jovanovich.
- Borkowicz, J. (1993). Portret niesentymalny. *Tygodnik Powszechny*, IV(9).
- Brinton, L. (1996). *Pragmatic Markers in English: grammaticalization and discourse functions*. Berlin; New York: Mouton de Gruyter.
- Brown, B. (1986). Cajun/English code-switching: A test of formal models. *Diversity and diachrony*, pages 399–406.
- Bulanin, L. (1970). *Фонетика современного русского языка*. Москва: Высшая школа.
- Chodakiewicz, M. (2012). *Intermarium: the land between the Black and Baltic Seas*. Transaction Publishers.

- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.
- Clyne, M. (1987). Constraints on code switching: how universal are they? *Linguistics*, 25:739–764.
- Clyne, M. (1991). *Community Languages: the Australian experience*. Cambridge, England; Melbourne: Cambridge University Press.
- Clyne, M. (2003). *Dynamics of Language Contact: English and immigrant languages*. Cambridge: Cambridge University Press.
- Daming, X., Xiaomei, W., and Wei, L. (2008). Social network analysis. In Wei, L. and Moyer, M., editors, *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*, pages 263–274. Blackwell Publishing.
- Davenport, M., D. M. and Hannahs, S. (2010). *Introducing Phonetics and Phonology*. Routledge.
- Dawlewicz, M. (2000). *Słownictwo socjolektu młodzieży polskiego pochodzenia w Wilnie (na podstawie badań ankietowych)*. Kielce: Elipsa.
- De Bot, K., Broersma, M., Isurin, L. (2009). Sources of triggering in code switching. In De Bot, K., Broersma, M., Isurin, L., editors, *Multidisciplinary Approaches to Code Switching*. John Benjamins Publishers.
- De Bot, K. (2000). A bilingual production model: Levelt's 'speaking' model adapted. *The Bilingualism Reader*, pages 420–442. Routledge.
- Deuchar, M. (2005). Congruence and Welsh-English code-switching. *Bilingualism: Language and Cognition*, 8(03):255–269.
- Dewaele, J.-M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development*, 25(2-3):204–222.

- Doğruöz, A. and Backus, A. (2009). Innovative constructions in Dutch Turkish: An assessment of ongoing contact-induced change. *Bilingualism: language and cognition*, 12(01):41–63.
- Doron, E. (1983). On a formal model of code-switching. In *Texas Linguistic Forum*, number 22, pages 35–59.
- Doroszewski, W., editor (2011). *Słownik języka polskiego*. Warszawa: Wydawnictwo Naukowe PWN.
- Dubisz, S., editor (2006). *Uniwersalny słownik języka polskiego*. Warszawa: Wydawnictwo Naukowe PWN.
- Dulewiczowa, I. (1993). *Gramatyka konfrontatywna rosyjsko-polska: fonetyka i fonologia, grafia i ortografia*. Warszawa: Sławistyczny Ośrodek Wydawniczy.
- Dwilewicz, B. (1997). *Język mieszkańców wsi Bujwidze na Wileńszczyźnie*. Warszawa: Wydawnictwo DiG.
- Edwards, M. and Dewaele, J.-M. (2007). Trilingual conversations: a window into multicompetence. *International Journal of Bilingualism*, 11(2):221–242.
- Ehala, M. and Zabrodskaia, A. (2013). Ethnolinguistic vitality and of ethnic groups in the Baltic countries. In Smyth, S. and Opitz, C., editors, *Negotiating Linguistic, Cultural and Social Identities in the Post-Soviet World*, pages 45–86. Bern: Peter Lang.
- Elistratov, V. S., editor (2002). *Словарь русского арго*. Москва: Русские словари.
- Fischer, K. (2006). Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume. In Fischer, K., editor, *Approaches to Discourse Particles*, pages 1–20. Oxford: Elsevier.
- Franceschini, R. (1998). Code-switching and the notion of code in linguistics: Proposals for a dual focus model. In Auer, P., editor, *Code-switching in*

- Conversation: language, interaction and identity*, pages 51–75. London: Routledge.
- Gafaranga, J. (2007). *Talk in Two Languages*. Palgrave Macmillan.
- Gardner-Chloros, P. (2009). *Code-Switching*. Cambridge: Cambridge University Press.
- Garrett, M. F. (1980). Levels of processing in sentence production. *Language production*, 1:177–220.
- Geben, K. (2001). Kalki leksykalne z języka litewskiego w języku polskim uczniów polskojęzycznych szkół na Wileńszczyźnie. *Poradnik Językowy*, (10):56–63.
- Geben, K. (2003). Wielowrazowe kalki leksykalne z języka rosyjskiego w języku uczniów polskojęzycznych szkół na Wileńszczyźnie. *Prace Filologiczne*, XLVIII:237–250.
- Geben, K. (2008). Język internautów wileńskich. *Poradnik Językowy*, (3):62–72.
- Gelles, R.J., Levine, A. (1999). *Sociology: An introduction*. Boston: McGraw-Hill College.
- Goffman, E. (1979). Footing. *Semiotica*, 25(1-2):1–30.
- Gorbachevich, K., editor (2004). *Большой академический словарь русского языка*. Санкт-Петербург: Институт лингвистических исследований (Российская академия наук).
- Guerini, F. (2006). *Language Alternation Strategies in Multilingual Settings: a case study – Ghanaian immigrants in Northern Italy*. Bern: Peter Lang.
- Gumperz, J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- Halliday, A. (2006). *Computational and Quantitative Studies. Collected Works of M.A.K. Halliday*. London: Continuum.

- Halmari, H. (1997). *Government and codeswitching: explaining American Finnish*, volume 12. John Benjamins Publishing.
- Hansen, K. (1987). Zum begriff der sprachsituation. *Ders.(Hg.), Studien zur Sprachvariation (unter besonderer Berücksichtigung des Englischen)*. Berlin: Humboldt-Universität, Sektion Anglistik/Amerikanistik, pages 6–49.
- Harrington, J. and Cassidy, S. (1999). *Techniques in Speech Acoustics*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Haspelmath, M. (2009). Lexical borrowing: concepts and issues. In Haspelmath, M. and Tadmor, U., editors, *Loanwords in the World's Languages: a comparative handbook*, pages 35–54. Berlin: Walter de Gruyter.
- Hogan-Brun, G. and Ramonienė, M. (2003). Emerging language and education policies in Lithuania. *Language Policy*, 2(1):27–45.
- Hogan-Brun, G. and Ramonienė, M. (2005). Perspectives on language attitudes and use in Lithuania's multilingual setting. *Journal of Multilingual and Multicultural Development*, 26(5):425–441.
- Hogan-Brun, G., Ramonienė, M. and Grumadienė, L. (2005). The language situation in Lithuania. *Journal of Baltic Studies*, 36(3):345–370.
- Holmes, J. (1988). 'Sort of' in New Zealand women's and men's speech. *Studia Linguistica*, 42(2):85–121.
- Hopper, P. and Traugott, E. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.
- Hudson, R. (1996). *Sociolinguistics*. Cambridge: Cambridge University Press.
- Hutchby, I. and Wooffitt, R. (2008). *Conversation Analysis*. Cambridge; Malden: Polity Press.

- Hymes, D. H. (1974). *Foundations in Sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press.
- Hymes, D. (1972). Models in the interaction of language and social life (revised version of Hymes (1967)). In Gumperz, J.J., Hymes, D., editors, *Directions in Sociolinguistics: The ethnography of communication*. Wiley-Blackwell.
- Jacobson, R. (1998). Conveying a broader message through bilingual discourse: An attempt at contrastive codeswitching research. In Jacobson, R., editor, *Codeswitching Worldwide, volume 1*, pages 51–76. Mouton de Gruyter.
- James, A. (1983). Compromisers in English: a cross-disciplinary approach to their interpersonal significance. *Journal of Pragmatics*, 7(2):191–206.
- Johanson, L. (2002). *Structural factors in Turkic language contacts*. Routledge.
- Joshi, A. (1985). Processing of sentences with intrasentential code switching. In Dowty, D.R., Karttunen, L. and Zwicky, A., editors, *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, pages 190–205. Cambridge University Press.
- Jucker, A. and Ziv, Y. (1998). Discourse markers: introduction. In Jucker, A. and Ziv, Y., editors, *Discourse Markers: descriptions and theory*, pages 1–12. John Benjamins Publishing.
- Juozeliūnienė, I. (1996). Individual and collective identity. In *Changes of Identity in Modern Lithuania. Social Studies*, pages 194–213. Vilnius: Lithuanian Institute of Philosophy and Sociology.
- Karaś, H. (2009). Kresowe cechy fonetyczne w polszczyźnie standardowej uczniów polskich szkół średnich na litwie. *LingVaria*, IV – 2(8).
- Karrebæk, M. (2003). Iconicity and structure in codeswitching. *International Journal of Bilingualism*, 7(4):407–441.

- Kawęcki, K. (2013). *Polacy na Wileńszczyźnie 1990-2012*. Warszawa: Europejskie Centrum Analiz Geopolitycznych Geopolitycznych.
- Klavans, J. (1985). The syntax of code-switching: Spanish and English. In *Proceedings of the Linguistic Symposium on Romance Languages*, volume 14, pages 213–31. John Benjamins Publishers.
- Krzywicki, T. (2005). *Litwa. Przewodnik*. Pruszków: Oficyna Wydawnicza "Rewasz".
- Kurcz, Z. (2005). *Mniejszość polska na Wileńszczyźnie: studium socjologiczne*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Kurzowa, Z. (1993). *Język polski Wileńszczyzny i Kresów północno-wschodnich XVI-XX*. Warszawa: Wydawnictwo Naukowe PWN.
- Kuzmich, V., editor (2000). *Жгучий глагол: словарь народной фразеологии*. Зелёный век.
- Lanza, E. and Svendsen, B. (2007). Tell me who your friends are and I might be able to tell you what language(s) you speak: social network analysis, multilingualism, and identity. *International Journal of Bilingualism*, 11(3):275–300.
- Levontina, I. and Shmelev, A. (2007). False emptiness: are so-called “parasitical words” really semantically void? *Meaning-Text Theory*, pages 259–268.
- Lipski, J. (2005). Code-switching or borrowing? No sé so no puedo decir, you know. In *Selected Proceedings of the Second Workshop on Spanish Sociolinguistics*, pages 1–15.
- Ljung, M. (2011). *Swearing: a cross-cultural linguistic study*. Palgrave Macmillan.
- Lüdi, G. and Py, B. (2009). To be or not to be. . . a plurilingual speaker. *International Journal of Multilingualism*, 6(2):154–167.
- Lyons, J. (1970). *New Horizons in Linguistics*. Harmondsworth: Penguin.

- Maschler, Y. (1994). Metalanguaging and discourse markers in bilingual conversation. *Language in Society*, 23(03):325–366.
- Maschler, Y. (1998). On the transition from code-switching to a mixed code. In Auer, P., editor, *Code-Switching in Conversation: language, interaction and identity*, pages 125–149. London: Routledge.
- Maschler, Y. (2009). *Metalanguage in Interaction: Hebrew discourse markers*. John Benjamins Publishing.
- Masojć, I. (1996). Czynniki kształtujące świadomość językową Polaków na Litwie po drugiej wojnie światowej. In Rieger, J., editor, *Język polski dawnych Kresów Wschodnich*, pages 25–45. Warszawa: Wydawnictwo Naukowe Semper.
- Matras, Y. (1998). Utterance modifiers and universals of grammatical borrowing. *Linguistics*, 36(2):281–332.
- Matras, Y. (2000). Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism*, 4(4):505–528.
- Matras, Y. (2009). *Language Contact*. Cambridge: Cambridge University Press.
- Matras, Y. and Sakel, J. (2007). Introduction. In Matras, Y. and Sakel, J., editors, *Grammatical Borrowing in Cross-Linguistic Perspective*, pages 1–14. Walter de Gruyter.
- McArthur, T. (1992). *The Oxford Guide to World English*. Oxford: Oxford University Press.
- Meeuwis, M. and Blommaert, J. (1994). The 'markedness model' and the absence of society: remarks on codeswitching. *Multilingua: Journal of Interlanguage Communication*, 13(4):387–423.
- Milroy, L. and Gordon, M. (2008). *Sociolinguistics: method and interpretation*. John Wiley & Sons.

- Milroy, L. and Wei, L. (1995). A social network approach to code-switching: the example of a bilingual community in Britain. In Milroy, L. and Muysken, P., editors, *One Speaker, Two Languages: cross-disciplinary perspectives on code-switching*, pages 136–157. Cambridge: Cambridge University Press.
- Myers-Scotton, C. (1993). *Social Motivations for Codeswitching: evidence from Africa*. Oxford: Clarendon Press.
- Myers-Scotton, C. (1995). A lexically based model of code-switching. In Milroy, L. and Muysken, P., editors, *One Speaker, Two Languages: cross-disciplinary perspectives on code-switching*, pages 233–256. Cambridge: Cambridge University Press.
- Myers-Scotton, C. (1997). *Duelling languages: Grammatical structure in codeswitching*. Oxford University Press.
- Myers-Scotton, C. (2002). *Contact Linguistics: bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, C. (2003). What lies beneath: Split (mixed) languages as contact phenomena. In Matras, Y., Bakker, P., editors, *The Mixed Language Debate: Theoretical and Empirical Advances*, pages 145 – 173. Mouton de Gruyter.
- Myers-Scotton, C. (2005). *Multiple Voices: an introduction to bilingualism*. Wiley-Blackwell.
- Myers-Scotton, C. and Bolonyai, A. (2001). Calculating speakers: codeswitching in a rational choice model. *Language in Society*, 30(1):1–28.
- Myers-Scotton, C. and Jake, J. (2000). Four types of morpheme: evidence from aphasia, code switching, and second-language acquisition. *Linguistics*, 38(6):1053–1100.

- Myers-Scotton, C. and Jake, J. (2001). Explaining aspects of codeswitching and their implications. In Nicol, J., editor, *One Mind, Two Languages: bilingual language processing*, pages 84–116. Blackwell Publishers.
- Nitsch, K. (1925). Język polski na Wileńszczyźnie. *Przegląd Współczesny*, XXII(33):25–32.
- Nortier, J. (1990). *Dutch-Moroccan Arabic Code Switching among Moroccans in the Netherlands*. Dordrecht, Holland: Foris Publications.
- Nortier, J. (2008). Types and sources of bilingual data. In Wei, L. and Moyer, M., editors, *Blackwell Guide to Research Methods in Bilingualism and Multilingualism*, pages 35–52. Blackwell Publishing.
- Offord, D. and Gogolitsyna, N. (2005). *Using Russian: a guide to contemporary usage*. Cambridge: Cambridge University Press.
- Plakans, A. (2011). *A concise history of the Baltic States*. Cambridge University Press.
- Poplack, S. (1980). Sometimes I'll start a sentence in English y termino en español. *Linguistics*, 18(7-8):581–618.
- Poplack, S. (2004). Code-switching. In Ammon, I.U., Dittmar, N., Mattheier, K. J. and Trudgill, P., editors, *Soziolinguistik. An International Handbook of the Science of Language*, pages 589–596. Berlin: Walter de Gruyter.
- Poplack, S. and Meechan, M. (1998). How languages fit together in codemixing. *International Journal of Bilingualism*, 2(2):127–38.
- Poplack, S., Sankoff, D. and Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26(1):47–104.
- Potašenko, G. (2007). *Lietuvos Tautinės Mažumos*. Kaunas: Šviesa.

- Ramonienė, M. and Extra, G. (2011). Multilingualism in Lithuanian cities: aims and outcomes of a home language survey in Vilnius, Kaunas and Klaipėda. *Kalbotyra: Germanų ir romany studijos*, (63):59–77.
- Riehl, C. (2005). Codeswitching in bilinguals: impacts of mental processes and language awareness. In *Proceedings of the Fourth International Symposium on Bilingualism*, pages 1945–1960.
- Sawicka, I. (1995). Fonologia. In Wróbel, H., editor, *Gramatyka współczesnego języka polskiego. Fonetyka i fonologia.*, pages 105–194. Warszawa: Wydawnictwo Instytutu Języka Polskiego PAN.
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.
- Scott, J. (2000). *Social Network Analysis: A Handbook*. SAGE Publications.
- Sebba, M. (1993). *London Jamaican: language systems in interaction*. London: Longman.
- Shvedova, N., editor (2005). *Русская грамматика*. Москва : Институт русского языка им. В.В. Виноградова.
- Silva-Corvalán, C. (1995). Lexico-syntactic modeling across the bilingual continuum. In Fisiak, J., editor, *Linguistic Change Under Contact Conditions*, pages 253–270. Berlin; New York: Mouton de Gruyter.
- Stapleton, K. (2010). Swearing. In Locher, M. and Graham, S., editors, *Interpersonal Pragmatics*, pages 289–306. Berlin: Walter de Gruyter.
- Stroud, C. (1998). Perspectives on cultural variability of discourse and some implications for code-switching. In *Code-switching in Conversation: language, interaction and identity*, pages 321–348. London: Routledge.
- Thomason, S. (2006). Language change and language contact. *Encyclopedia of Language & Linguistics*, 2:339–46. Elsevier.

- Treffers-Daller, J. (1991). Towards a uniform approach to code-switching and borrowing. *ESF Network on Codeswitching and Language Contact. Papers for the Workshop on Constraints, Conditions and Models*, pages 259–279.
- Turska, H. (1982). *O powstaniu polskich obszarów językowych na Wileńszczyźnie*. Wrocław: Wydawnictwo Ossolińskie.
- Valdés, G. (1981). Code-switching as a deliberate verbal strategy: a microanalysis of direct and indirect requests among bilingual chicano speakers. In Duran, R., editor, *Latino Language and Communicative Behavior*, pages 95–108. Norwood, N.J.: Ablex Publishing Corporation.
- Van Gijn, R. (2009). The phonology of mixed languages. *Journal of Pidgin and Creole Languages*, 24(1):91–117.
- Van Wijnendaele, I. and Brysbaert, M. (2002). Visual word recognition in bilinguals: phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance*, 28(3):616.
- Wajszczuk, J. (2005). *O metatekście*. Warszawa: Katedra Lingwistyki Formalnej Uniwersytetu Warszawskiego.
- Walters, J. (2014). *Bilingualism: the sociopragmatic-psycholinguistic interface*. Routledge.
- Warner, N. (2012). Methods for studying spontaneous speech. In Cohn, A., Fougeron, C. and Huffman, M., editors, *The Oxford Handbook of Laboratory Phonology*, pages 612–633. Oxford: Oxford University Press.
- Wei, L. (1994). *Three Generations, Two Languages, One Family: language choice and language shift in a Chinese community in Britain*. Clevedon: Multilingual Matters.

- Wei, L. (1998). The 'why' and 'how' questions in the analysis of conversational code-switching. In Auer, P., editor, *Code-switching in Conversation: language, interaction and identity*, pages 156–179. London: Routledge.
- Wei, L. (2005). How can you tell? Towards a common sense explanation of conversational code-switching. *Journal of Pragmatics*, 37(3):375–389.
- Wei, L. (2008). Research perspectives on bilingualism and multilingualism. In Wei, L. and Moyer, M., editors, *Blackwell Guide to Research Methods in Bilingualism and Multilingualism*, pages 3–17. Blackwell Publishing.
- Wiebusch, T. and Tadmor, U. (2009). Loanwords in Mandarin Chinese. In Haspelmath, M. and Tadmor, U., editors, *Loanwords in the World's Languages: a comparative handbook*, pages 575–598. Berlin: Walter de Gruyter.
- Wolf, H.-G. (2001). *English in Cameroon*, volume 85. Walter de Gruyter.
- Zemskaya, E. (1979). *Русская разговорная речь: лингвистический анализ и проблемы обучения*. Москва: Русский язык.
- Zentella, A. (1997). *Growing up Bilingual: Puerto Rican children in New York*. Malden: Blackwell Publishers.
- Ziem, A. (2014). *Frames of Understanding in Text and Discourse: Theoretical foundations and descriptive applications*, volume 48. John Benjamins Publishing Company.

Appendix A

Data Collection Forms

A.1 Letter to participants



Brigita Brazyte

Centre for Linguistics and Philology

Walton Street

Oxford OX1 2HG

E-mail: brigita.brazyte@ling-phil.ox.ac.uk

[English version]

In partnership with the University of Oxford (UK), we are currently recruiting participants for a linguistic project in order to investigate the language use in the Polish community in Vilnius, Lithuania. Please read carefully through the information provided below, and if you meet the required criteria, we strongly encourage you to take part.

What are we trying to find out?

We aim to investigate the consequences of the co-existence of Polish, Russian and Lithuanian in the everyday language practices of the members of the Polish community in Lithuania. What impact has this had on the local variety of Polish? Is there any overlap between the three languages?

What are the selection criteria?

We are looking for members of the Polish community in Vilnius who:

- are between 20 and 30 years old;
- speak Polish, Russian and/or Lithuanian on a regular basis.

What will happen if you take part?

You will be given a digital recorder and asked to record a casual natural conversation between you and your one or two friends. The recording should be between 60 and 90 minutes long. You may choose a time and place that is convenient to you both. Before you start, you and your friend(s) will be given an opportunity to ask any

questions. Once you are confident about what you are expected to do, you will be given a brief questionnaire to fill in and a form to sign, confirming your voluntary participation.

During the meeting, you and the researcher will decide on how many days you may need to do the recording and agree on a date when you will return the recorder. You are welcome to use your own software and forward the mp3 of the recording afterwards.

What happens with the recordings and questionnaires?

The questionnaires will serve to gather some basic information about the linguistic profile of each participant. They will not be discussed in any subsequent analysis and will be conducted merely to ensure that the basic selection criteria have been met.

The recordings will be transcribed, codified and anonymised, to make sure that any identifiable personal information is not featured. The identities of each participant revealed in the questionnaires and recordings will be kept strictly confidential. Some extracts from the recordings will feature as examples in the written analysis or oral presentations that will result from the recordings.

Who is conducting this research?

The research project is organised by Brigita Brazyste, who is a graduate (doctoral) student under the supervision of Dr Jan Fellerer. She will be the researcher who is in charge of arranging the recordings and answering any questions that you may have.

What should I do next?

If you would like to take part in this study or to discuss the research with someone beforehand, please feel free to e-mail brigita.brazyste@ling-phil.ox.ac.uk or use the contact details in the letterhead.

A.2 Consent form



Brigita Brazyte

Centre for Linguistics and Philology

Walton Street

Oxford OX1 2HG

E-mail: brigita.brazyte@ling-phil.ox.ac.uk

Telephone number: 0044 752 343 4080

[English version]

The Patterns of Language Use in the Polish Community in Lithuania

Thank you for expressing your interest in participating in this project run by a doctoral student of Oxford University, which aims to investigate the patterns of language use in the Polish Community in Lithuania.

If you decide to take part, you will be asked to fill in a brief questionnaire about the languages that you use in your day-to-day activities and record a conversation between you and your friend(s), between 60 and 90 minutes long.

If you agree to take part, please fill in the form below and return it to the researcher in charge as soon as possible.

To find out more about the project, please read the attached information sheet or e-mail me at brigita.brazyte@ling-phil.ox.ac.uk.

Name: Date of birth:

I have read and understood the details of the above study, and have had the opportunity to ask questions and discuss it with others. I understand that participation is voluntary and that I am free to withdraw at any time without giving any reason. I agree to take part in the above study.

Name: Signature:

Date:

A.3 Questionnaire



Brigita Brazyte

Centre for Linguistics and Philology

Walton Street

Oxford OX1 2HG

E-mail: brigita.brazyte@ling-phil.ox.ac.uk

[English version]

Why a questionnaire?

To make sure that all the participants fulfil the basic selection criteria, as well as to put our results into context, we need some background information about the participants. All information provided will be entirely anonymous, and only related to the equally anonymised recordings.

Questionnaire

Year of birth:

How would you describe your ethnic background?

Please rate your fluency in each of these languages on a scale from 1 ('limited command') to 5 ('fluent'):

1. local variety of Polish
2. Standard Polish
3. Russian
4. Lithuanian

What language do you consider to be your native language?

What language(s) do you speak at home?

What language(s) do you speak at work?

What language(s) do you speak with your peers?

Appendix B

Transcription Extract

<File name = 120919/001.mp3>
<Setting = At home. Speakers looking at photos.>
<Speaker S = Male, 22 years old>
<Speaker T = Male, 23 years old>
<Recorded = September 2012>
<Transcribed = Brigita Brazyte May 2012>
<Notes: conversation about cars>

- S: a **prikin'** my tak **vo**— tego— godzina przegaliby a wtedy **ponjal** że nie wkluczyli
- T: ((laugher)) **i takoe byvaet**
- S: ty co patrzysz?
- T: [LT] **Eurostandartas** trzeba nowiejsze patrzeć
- S: Eu— Eu— Euro trzy— cztery lepiej bierz
- T: **nu da**
- S: **pilotka** jeszcze **ne vsem** ()
- T: jeszcze **kolxoz** u nas
- S: to **vot** jaka () kolejka (czterdziesta)
- T: i na jakiej ty chcesz jeździć?
- S: jeżeli **čestno** ja to **libo** na Volve **libo** na Mersie

- T: a ja na Dafie chciałby (.) ale Scania chcem popaczyć (.) **nu vo** sto pięćdziesiąt trzy tysięcy
- S: pamiętam **bat'ka** jeszcze jeździł na Scanie (.) **bljaxa udobna mašina**/maszyna
- T: ale co ona stara (.) siódmego roku
- S: **nu** (.) ale tak—
- T: kulega **vot** dwa lata **nazad** poszedł pracować (.) on dostał (.) sama nowa
- S: którego?
- T: **nu** on jedenastego roku poszedł pracować jedenastego i dostał (.) Mersa (.) (Hectasa)
- S: aj wiem (.) u **bat'ki** taki że
- T: mhm?
- S: u **bat'ki** mojego taki że samy (.) jedenastego roku
- T: jemu **vot** niedawno przypisali dziesięć tysięcy (.) dziesięć tysięcy litów (.) **nu jak tipa**
- S: zarobił **tipa** czy?
- T: mhhh
- S: czy **avansom?** czy prosto sprzedał **fura** (.) **po ǰernomu?**
- T: nie (.) on przyjechał **tipa** [LT] **priekaba** patrzy że u jego bendzi jeżeli— **tipa** pubita (.) z tyłu (.) a nie on pobił ni co (.) prosto
- S: a (.) **pripisali** tego—
- T: **pripisali** —
- S: **kozel otpušenija**
- T: bo— **nu— nu da** (.) **tipa** w tym **smysle** bo przy tym co jechał on nie popatrzył (.) **nu** to jego **ošibka**
- S: **nu da** (.) to też prawda
- T: **nu** on tam ni wiem (.) on ta na to kręci żeby tym co jeździł ten wypłacił a on— **dokazatel'stva** on ni ma
- S: **nu da** (.) ()
- T: ja to ni wiem (.) (zasiąd) myśla jakiścik wzięści czy **kamera** czy coś takiego kupić sobie żeby **čisto** —
- S: a to przed **rejsom**—

T: nie **nu** potem—

S: **kame**— **kamery vokrug fury oboideš'** bendzisz patrzeć

T: nie (.) **polnost'ju** i [LT] **krovimas** żeby **vse** filmować (.) **nu** tak (.) żeby było

S: to trzeba (.) **video registrator** postawić

T: **video registrator** jaki (.) że— **nu** byłoby prosto—

S: **dokazatel'stvo**

T: **dokazatel'stvo** (.) **da**

S: **nu** (.) bendzisz zlewać (.) to zobaczo

T: **nu tak da** (.) i zrazu **bačok** sfilmujesz (.) ile zalał ((laughter))

S: ((laughter)) powiedzo ty zlał trzydzieści (.) nie (.) dwadzieścia

T: **da** (.) tylko nie stukaj **v pustuju**

S: a (.) co za **telefon** teraz?

T: czekaj (.) **ne ponjal ja**

S: czego?

T: (0.3) zlewaniu (.) **glavnoe** o czym o czym o zlewaniu (.) (ale) czekaj (.) a gdzie tu [LT] **gultas?** ta **vo** musi u tego

S: tak (.) ta zdejmują się (.) **vo** tu **vo** z dołu jeszcze jedna jest (.) **tipa** w dole **spal'nik**

T: aha

S: on— on tak odchyla się

T: a odchyla się i **vse** (.) i **tipa** śpisz?

S: **da** (.) i tu nie pa— nie pa—

T: franki zawiesić

S: **da** (.) ja nie pamiętam jego **libo**— **libo** bendo żelazne te **tipa opory** (.) **nu tipa** te (.) re— ry— rączki

T: **nu tipa** jak **na amortizatorax** jak spuska się **da**?

S: **ne amortizatory** (.) jak **gidravlika tipa** (.) jak **gidravlika na eto vot** (.) **kapot** który podejmują

T: **nu ja ponjal**

S: **nu** (.) ale (.) latem na wierzchu nikiedy ni śpi

T: a tam dwie miejsca?

S: **v smysle** dwie miejsca?

T: **nu** dla dwóch

S: **da** (.) jeden z dołu drugi z wierzchu

T: aha

S: ale na wierzchu latem nikiedy ni śpi (.) zgutujisz się

T: **nu tak**

S: odkry— odkrywasz **ljuk** (.) okny troche i na doł (.) i **nadejsja** że wyżyjisz

T: ja myśla (.) jeszcze w Italii / **v Italii** jakie (.) czy **v Ispanii**

S: to **vo** swego— z **bat'ko v Italii** stoim je— jeden jego ten znjomy u nas z roboty (.) sio— do jego zaszli (.) **mašina**/maszyna zaprowadził jego **nu tipa kondiška** wkluczył (.) i dobrze (.) zimno (.) na ulica tylko wychodzisz i zrazu gu— **vse** ciplej i ciplej i ciplej i ciplej

T: mhm

S: i tylko lecisz do **mašiny**/maszyny żeby i znów **kondiška** wkluczyć

T: ni można tak (.) zachorujisz zrazu

S: **nu** (.) ależ jak gorąco to też ni fajnie

T: **vo** (.) coś podobnego

S: **mexa— mexanika** ()

T: tutaj (trip tronik)

S: m— aj (.) sam sobie możesz to **vo**

T: **poluavtomat**

S: chcesz? ((offering something))

T: ni odkażam się

S: **na xaljavu i uksus sladkij**

T: to prawda

S: i jeszcze jaka (.) **a druguju skanejku vkluči**

T: czekaj (.) **skanejku?** jak—

S: mhm (.) Skania (.) **skanejka**

T: przy— przy— **tipa** ta (.) (2.0) eh— (.) przezywają Iveko **ovečki**

- S: Iveko **čudo sveta dlja sovremennogo čeloveka**
- T: Vadim gadał że tam **izoljacija** bardzo drena
- S: a jakie zdjęcia malutkie
- T: mówi (.) **tipa** przed spaniem tam wkluczył na przykład **esli** gazowa ta (.) pieczka wkluczył (.) i:: (.) potem ona powietrze bardzo zabiera
- S: mhm
- T: troche poleżał i **vse** (.) i zimno
- S: jeden opowiadał (.) eh (.) Nowy Rok spotykał **v Italii**
- T: mhm
- S: **bljaxa** wiesz a czo— a w nocy tam zimno bywa (.) a on tylko n— **nu** (.) le— letnie udzienia (.) wził swój **gazovyj**—
- T: gdzie spotykał (.) **v Italii**?
- S: tak (.) **na severe** (.) Nowy Rok spotykał koło **avstrijskoj granicy**
- T: mhm
- S: wiesz **koroče** mówi patrza zimno **blja** w nocy spać (.) wził wkluczył **nu** tam **svoj ga— gazova** ta (.) **nu** ten **ballon**
- T: mhm
- S: wkluczył jego—
- T: on a powietrze zebiera bardzo
- S: **da** (.) **no** wiesz koło jego śpiz to **pofig** tobie (.) **xot'** cieplej (.) a jego te— te **nejlonova** ta była (.) **nu** (.) koldra
- T: mhm
- S: zwała się i pa— na gaz patrza —
- T: zapaliła się?
- S: — **blja** ten naczał ręko i tuszyć (.) wszystko— wszystka ręka **v plastmasse**
- T: **užasno**
- S: **nu**
- T: nie (.) te **avtonomki** musi sama lepiej sa— **dizel'nyje**
- S: **da** (.) wkluczył zapomniał
- T: **da** (.) tak ja jak **von** jak pracował tak mnie nerwowali te ichne **avtonomi**

- S: **če tak?**
- T: sam chcesz troche pospać (.) tam w tej budce (.) a tutaj **fura** specjalnie podpedziwszy bliżej wiesz on płaci tak że wiesz—
- S: mhh
- T: — tego (.) to (.) tylko— tylko tak śpisz śpisz (.) tylko jak wkluczą podskoczysz zrazu wiesz **fura** co takiego (.) szuch szuch szuch (.) **vse** normalnie dalej (.) (Polak) był przyjechał ten—
- S: **nu** (.) pamiętam patrzy pracował na tej (.) **na mojkę** (.) i **vot nu** taka **skanejka** do nas przyjechała trzeba było głowa podjąć żeby motor pomyć
- T: **nu**
- S: podejmujim (.) podejmujim (.) podejmujim (.) a tam żesz jest ta **nu** (.) hydraulika która (jej) podtrzymuji
- T: **nu**
- S: a patrzym a on a tam złamana była (.) i głowa jak **pošla** tak i **pošla**
- T: zwała się?
- S: ni zwała się (.) rękami wstrzymali (.) ta głowa ja chciał powiesić—
- T: a co ona ni ciężka? czy duża prosto?
- S: my dz— sześciu byli
- T: a:::
- S: **nu** on a tam jeszcze dobrze **zavesy** jeszcze byli dobre (.) my tak wstrzymali i **nazad** () zrzucili (.) (powiedzieli) **idi ty nafig** jej myć (.) () z tym **motorom**
- T: **da blin** (.) **tut**— (.) **nu** ja podejmowawszy w tej (.) w Iveco (.) w cargo tej co **vo**
- S: aj (.) jakoby uczyli się
- T: **nu** jak uczyli się (.) drugie tam nikt nie podejmował (.) to ja prosto podeszedł tam pomogci to dla mnie i dał **xaltura** pomogci z emblemy te stoi
- S: **vot** () () **čisto kranik** stoi i:: (.) na Ivece nie pamiętam gdzie stoi na trąbka (.) a w Scanie pod tym **vo kapotom** podejmujisz **nu** ta (.) i ona tak **vo pokrečka** z **emblemoj** tam stoi (.) wybierasz
- T: mhm (.) tam— tam z lewej strony bierzysz ta sama— dwie **palki v salone** (.) wyciągujisz (.) tam takie bierzysz przykręcasz i potem kaczasz

S: jak (Tamkatem)

T: **da** (.) **nu** jeszcze trzeba odsunąć tam ten troche

S: jeżeli bendzisz ni odkryysz tego to rozłamisiz prosto

T: co?

S: **nu**

T: **kapot** trzeba odkrywać

S: **da**

T: **nu da**

S: **Mersa vključi**

T: co (.) ładne te (.) raniej mnie ni podobali sie (.) a teraz wiesz jak zaciekawiony jestem to

S: ((whispreing)) Mercedes Mers

T: **dopustim** aby jaka

S: (0.3) a to (formulé)

T: **nu** to ile (.) **vo** (.) stary

S: sta:::ry już (.) **vo** wiesz co dwieście czterdzieści jeden znaczy?

T: co?

S: czterdzieści jeden znaczy?

T: **tipa** [LT] pajégumas jakiś

S: mhh?

T: [LT] pajégumas jakiś czy co

S: mhm (.) czterysta **desjat' lošadej** (.) a osiemnaście co znaczy?

T: co ja wiem

S: **model' kuzova**

T: mhm (.) co to (.) **vo** (.) **devjanosto vos'mogo** (.) **nu** i na takiej gdzie ty wyjedzisz ty w/v **rejs**

S: a::: ciebie nigdzie ni puszczo (.) () zapłacisz gdzie ni puszczo

T: **nu tak** tak (.) dlatego oni i stojo i sprzedajo się dlatego

S: to::: wiesz kto ich skupuji? **Kazaxstan**

- T: **nu tak** oni **vse** bioro (.) **Kazaxi** tam (.) **mašiny** te po tysiąc lity— po tysiąc eury tam jakie (.) skupujo (.) i co oni z tym zarabiajo?
- S: jeżdżo na ich (.) myślisz ich drogi— **nu** (.) na ich drogach pojeździsz normalnie
- T: **nu vot**— ((coughing))
- S: o (.) **vo** te **vo** pokaż
- T: ten?
- S: nie (.) niżej (.) ten
- T: (0.3) jakiś on—
- S: **bljaxa** cze— chło—
- T: **bampera kak budto** ni ma
- S: i **bampera** ni ma (.) i koła za malutkie (.) patrzaj (.) widzisz jak on podjęty wysoko?
- T: **nu** (.) coś tu z nim **tut** narobione
- S: coś— a:: musi— **karjernik** musi
- T: może (0.3) **da** (.) **tut on/on**
- S: () jakaś tam (.) nie widziawszy
- T: Polaki (.) co ty chcesz
- S: mh (.) t— to samo co u **bat’ki**
- T: **v salone ničego**
- S: aha (.) **vo tut koroče pacan** który (()) ta całej konsoli—
- T: mhm
- S: — **nu** (.) jest **stupen’ka** (.) naci— naciskujesz byle czym na— na spanie wiesz rozebrać się
- T: mhm
- S: — dwa **bolta** odkry— odkręcasz po boku (.) i ta **plastmaska** wybiera się ((laughter))
- T: **nu ja ponjal** ciebie
- S: **da da** (.) tak samo **pol avtomat**
- T: **da**

- S: **nu** ((laughter)) a poczekaj (.) jeszcze przekręć (.) jiszczé (.) jiszczé (.) **do salona dojdi** (0.2) **ešče** (.) **ešče** (.) **vo** (.) to **vo** całe wybiera się i włązi **gde to** pięć
- T: **da** (.) **tak tut i vse**—
- S: ja widział w Skaniejkach co można— wiele można znachodź
- T: **vsego**
- S: to już dwunasty rok
- T: (0.3) ((laughter)) **da** (.) poczekaj (.) to motor (.) nie (.) to k—
- S: to **korobka**
- T: **da** (.) widza (0.2) o (.) tu już ja/ **ja ponimaju** pokazane jak spać **vse**
- S: **nu** (.) i tu też można spać
- T: a telewizor **vse** trzeba swój ciągnąć jakiś tam (.) musi (.) nie bywa tam w tych
- S: nie (.) **vot tut vo vo** na głowa (.) poczekaj (.) **vot** tego— w tym **vo** miejscu—
- T: mhm
- S: — ta **tipa** jak— jak **bardačok** taki **zdorovyj**—
- T: **nu**
- S: — to tam **televizor** letko postawisz jeżeli co
- T: **nu** ale trzeba jak (.) DVD jaki tam
- S: a co ja wiem (.) wziął normalnie **telik**— **nu tipa** telewizor gdzie z boku **disk** wkładasz i **vse** (.) a jeżeli tobie potrzebny **kompjuter s navigatorom** to z **loxtopom jezdíš'**/ jeździsz
- T: sama lepiej musi **leptop** z sobo mieć i **vse**
- S: mhh (.) jak komu
- T: wziąć i przykręcić gdzie żeby nie skradliby
- S: (0.3) **nu** (.) brat bierzy to **v Ispanii** gdzie/ **gde vaj faj na xaljavu na zapravkax** (.) hop podklużył się i dla żony dzwoni (.) **po Skajpu**
- T: **nu tak a: da** (.) tak (a) popróbuj dzwonić tak
- S: **nu** (.) jeden raz podzwon— **nu** jemy żona jeden raz podzwoniła (.) **rešil aj pofig** pogadam
- T: mhm

- S: on pierwszy raz **bljad'** wyjechał (.) nagadał **gde-to na** pięćset— na pięćset lity za pare dni (.) jemu to i wyliczyli **s polučki** ((laughter))
- T: aj (.) jeszcze musi na roboczy(m)
- S: tak (.) (u) jego Tele 2
- T: mhh (.) **da::** (0.4) **ničego** taki
- S: fajny (.) **aparatik**
- T: (0.2) aj (ja) ni wiem (.) ja jak pomyśla czym dalej tym straszniej to **vse** dla mnie
- S: co? jeździć? ()
- T: z to jazdo
- S: czego tak? tak? (0.2) co tu straszego?
- T: wielka odpowiedzialność **vse**
- S: a gdzie jej ni ma?
- T: **nu v smysle—**
- S: **konečno— onečno ty jak rabotal storožem—**
- T: mhm
- S: — kolegów nawołał (.) napilsia i **vse**
- T: mhh
- S: (0.7) kolor mnie nie bardzo
- T: a jaka różnica jaki tam kolor (.) będziesz tam—
- S: mh (.) wiesz jak raniej— po staremu robili? o ta **vo plastmassa—**
- T: co ty na niej będziesz do panienek jeździć **čto li?**
- S: to **vo** (.) **vo** te **plastmassovyje** podejmowali i u nich była w środku jeszcze **niša**
- T: mhm
- S: **tipa nu** (.) **dvojnaja stenka** (.) środek pusty (.) to **vo** z dołu przekrajali (.) dziura i (.) pszyk pszyk pszyk
- T: mhm
- S: i nakładali
- T: przekrajali wszystko a potem jak [LT] **darbdavys** na przykład—

S: myślisz [LT] **darbdavys** nie wiedział?

T: **nu tak** myśla wiedział (.) dzielili się z nim (0.2) samochód testowy (0.2) to on był na [ENG] test drive (pochoda) (.) z **salon**-u jakiego to

S: mhm

T: i **vse** i— a sprzedają się **vse** wy— posprzedawali i ten sprzedają (.) na test drive (.) u jego już i **probeg i vse** jest (.) na nim katano się

S: mhm

T: taniej sprzedają

S: (0.6) mm (.) co jeszcze? dawaj Dach popatrzym

T: dawaj (.) popatrzym Dachy (0.3) (Daewu)

S: ((talking while laughing)) (deutsch) **fura**

T: (dodży) (.) ni będzi (.) **Amerikancev** (0.3) ni— myśla już **vse** już (.) latem iści pracować(.) zrazu (.) ni wiem czy zagranica jechać ale tak

S: **tipa** po Litwie pokatać się żeby **privyknut**'?

T: **nu da::** (.) żeby wiesz—

S: hop (.) **v Klajpedu sgonjal i vse**

T: mhh (.) do Iki jakiej iści **von** na ta

S: **na kubik**?

T: nie (.) widział w Iki jest i ten (.) **fury** normalne

S: a **kubikom** czemu ni chcesz?

T: **nu** czy **na kubik** (.) jaka różnica (.) nie (.) żeby **tipa pricep** jeszcze byłby (.) przyczepiony

S: **da?** będzi

T: mhm (.) nie **nu** z jednej strony **esli** u mnie będzi C i E ja przyjda do tego to mnie i dadzo na ten jechać (.) **nu esli** — **esli** dadzo jeszcze **konečno**

S: może dadzo

T: mhm

S: tam **po idei** nawet w Maximie jeżeli pracujisz (0.3)—

T: mhh

S: tam jeżeli dobrze **vse** to w miesiąc trz— trzy dostajisz

T: trzy?

- S: trzy (.) brata znajomy pracuji (0.2) ale powiedział te— tem— **nu::** (.) z ra— **nu** (.) kaž— każda noc w chacie nocujisz
- T: mhm
- S: **nu tam** (.) **po**/po małym **isključenijam**
- T: **nu esli** już tam coś stani się
- S: **nu da**
- T: — z maszyno czy co
- S: **nu dopustim v Klajpedu poexal** nie zdążył (.) nie mogą **tipa** nie mogo **tipa** rozgruzić już
- T: tak ale—
- S: **nu i vse** musisz— (.) to trzeba zostawać się
- T: ale do Klajpedy **nu vse ravno** patrzaj mnie— co z tym **otstoem** (.) (bardzo) ciekawie że **vse ravno** do Klajpedy tej **devjanosto** pojedzisz jakich tam **dopustim** (.) dobrze (.) wszędzie magistrale tam jedzi wszystko (.) ile—
- S: osiemdziesiąt licz zawsze
- T: **nu** osiemdziesiąt (.) **nu** pojedzisz (.) **vse** (.) i cztery z połowo godziny przejedzisz ty musisz stanąć (.) to gdzie ty tylko za cztery z połowo godziny **dopustim** ty dojedzisz tam
- S: ((sigh)) (0.4) do Kaunasa dojedzisz zz— już **sdelaes'** — do Kaunasa godzina jechać za Kaunasem przejedzisz troche **delaes'** **otstoj** (.) za Kaunasem licz so— so— **sotku delaes' i vse**
- T: mhm
- S: **otstoj delaes' i dopiliš'** (.) dojedzisz (.) ciebie rozgrurzo (.) **otstoj avtomatom i** (.) ((funny noise)) **nazad** (.) ty że sam pojedzisz?
- T: **nu**—
- S: ale co minus to mówio że ty musisz **pilit'**— **sčitaj bez otstoev ty tam jezdiš'** (.) jedź **i vse**
- T: **nu tak a: a:** płacić?
- S: **tipa štrafy?**
- T: **nu**
- S: **tvoi problemy**
- T: to w tym **delo** że—

S: **nu** (0.3) ()

T: ni wiem starać się **prosto** n— (.) sama gorzej że—

S: słuchaj (.) a on zapisu—? a (.) on zapisuji

T: zapisuji (.) zapisuji (.) ty wcale w technice nie rozbierasz się tylko jak gdzie co schować i (0.1) złać

S: ej!

T: ((laughter))

S: trzeba gadać w imię rady

T: **nu da**

S: ty mnie jeszcze bendzisz— w nocy bendzisz mnie dzwonić (.) w Litwa tylko wjedzisz (.) a jak to trzeba zrobić (.) jak złać (.) gdzie schować

T: ty bendzisz mnie jeszcze na granicy stać czekać

S: mhm (.) s **otkrytymi ob"jat'jami i pustym bakom**

T: **da**

S: ((inaudible))

T: **da** (.) wielki

S: czekaj (.) wiele to (.) centymetr to wiele to litry ja zapomniał

T: **rgde-to sto**/sto

S: centymeter?

T: **vrode by da** (0.1) ile mnie po—

S: trzy ba— trzy baki normalne to **eto poltoraška** (.) numer **sem'** ((yawning)) poczekaj **vo** (.) wklucz **nazad** (.) **vo** numer **sem' tut vo** (.) jak **ljuk** jest tak—

T: mhm

S: — to jest tak **tipa nu** jakby **perila** (.) **dopustim nu** chcesz ręcznik powiesić (.) **bat'ka** tam kielbasa wiesił ((laughter)) suszyć ((laughter))

T: **da::**

S: **nu** (.) **vo tut lestnica** spuska się u ciebie (.) tych (.) **vskarabkalsja**

T: fajnie

- S: w Skaniejke starej to wiesz jak robi się? ona **vo tak vo nu** ta (.) **nu** (.) drabina **vo tak vo** wychilona (.) **nu vo tak vo** (.) spuskasz się i ty po jej włazisz na (ten) drugi **etaž**
- T: jak w domu (.) w (Dafach) **vrode** liczy się sama wielka **kabina**
- S: mhm (.) ()
- T: ja by i chciał czym większa tym lepiej (0.2) tym **prostornee**
- S: mhm (.) jest— tylko ni bierz jeżeli bendzi Skaniejka tobie dadzo (.) mogo dać bez tej **šljapy**
- T: mhm
- S: () siedzieć to **pofig** ale
- T: nie **nu tak** ty **esli** już jedziesz to ty rozpatrujisz i te samo— ichne maszyny i **vse** (.) ni tak że przyjdiesz **vse** mnie ta robotka bardzo potrzebna (.) dajcie i **vse**
- S: he he
- T: oni ciebie posadzo niewiadomo na co
- S: tak i bendzi
- T: mhh (.) (**nu**) z jednej strony dla młodego to tam
- S: ty idzisz— ty **nikto** (.) **iz ulicy kak**
- T: mhh?
- S: przyjdiesz (.) co dadzo na tym pojedziesz
- T: **da**
- S: (co) palce **pazdvineš'** i powiesz mnie trzeba sama **pontova fura**
- T: **da** (.) **dajte mne**
- S: aha (.) i **polučka desjat' štuk** (.) **avansom vpered**
- T: **da**
- S: i **budeš' pit' na mesjac** (.) powiesz **ladno** (.) **mogu exat' uže**
- T: **da::** (.) **na mesjac desjat' štuk da?**
- S: **nu** (.) na miesiąc? tydzień
- T: wielkie u ciebie **potrebnosti**
- S: **ne u menja takije potrebnosti** (.) dadzo **neudobna fura** i powiedzo **ežžaj v Kazaxstan**

T: mhm

S: przyjedziesz cały złamany

T: ja na/ **na gol'fe** ile odjeździł to już to wiesz

S: co? **v Kazaxstan?**

T: **nu** ja po Litwie po osiemset kilometry robił

S: **na swoim gol'fe?**

T: **da** (0.6) znowu bendzim patrzeć

S: same drogi patrząj

T: sama nowa (.) sama droga (.) sama dobra

S: (0.6) ((yawning)) nam nadojadło gadać to jak nie nadoje słuchać

T: ((laughter)) widzisz (.) **tut** ni ma tego (.) **kozyrka**

S: a **nafiga** on tobie?

T: ona i tak **po idei** musi bendzi wielka

S: ona wielka (.) tobie starczy (0.2) DAF (.) to na wszystkie DAF-y **avtomaty każe't'sja** ()

T: widzisz jaki **divan** (.) **udobny(j)** (0.3) **ničego**

S: **bak zdorovyj**

T: **da** (.) **bak** tu **zdorovyj** (.) i ni jeden (.) z drugiej strony

S: z drugiej strony () krótszy

T: mhm (.) **nu tak** dla ich ile—

S: ich **voobše** ni ma

T: mh?

S: ich **voobše** ni ma (.) tam/**tam** jeden **bak**

T: mhh

S: ile tam włazi?

T: mhh?

S: ile tam włazi? ja zapomniał

T: a kiedy ty zbierasz się iść uczyć się już

S: ja na jasięń benda

T: kiedy to bendzi?

S: szesnastego listopada

T: mhm (.) tak (.) tobie wychodzi—

S: teraz dziewiętnaście (.) () **matematika**

T: to jeszcze (.) więcej roku (.) (trzeba) czekać (.) to **posle** szkoły

S: prze- przed świętami

T: mhm (.) na drugi rok przed świętami

S: na ten rok

T: w tym roku u ciebie dwadzieścia bendzi?

S: w tym roku dwadzieścia bendzi

T: aa (.) ja myślał że tylko tobie dziewiętnaście w tym roku—

S: nie::

T: — **spolnitsja**

S: **ne obizaj e moe (.) ja znaju čto ja molodo vygljažu no ne nastol'ko (.) te vo** jeszcze—

T: to jeszcze nie strasznie

S: jeszcze przekręć (.) do przodu (.) **eše (.) vo (.) koroče** jak brat mnie opowiadał (.) na tym jeździ

T: on na DAFie jeździ?

S: **da (.) napelnia (.)** było już **koroče pod golovinoj (.) poduški** podejmują i zapełniają (.) jeszcze **gde-to** dwadzieścia litry wchodzi które on potem zlewa

T: poduszki podejmują?

S: **da (.) zadnij most** podejmują (.) tył— tył podejmują

T: a podejmują się on **da**?

S: **da (.)** czy— nie pamiętam czy on podejmują tyłem ja nie pamiętam czy **to li dupo— nu (.)** przód opuska żeby tył pojeźdź

T: aha

S: jakościk tak robi

T: i/ **i fury daže** zapełniają jeszcze więcej (0.2) **vse** można (.) Litwini **vse** przemyśle

S: to **Russkije vse** przemyśleli

- T: mhh (.) ni wiem (.) jak nazywa się **pod naklonom** (.) jak na rynku byli w (.) przeszły (.) tydzień (.) **nu** w subota jak mówił że był na rynku to (.) też tam dzwonił m- maszyna szukają żeby były **bak** wielki (.) **nu** tam potrzebny dla/ **dlja dela** i (.) ehh (.) ta **tipa** (.) tam dzwonia pytamsia jak tam tego (.) czy— ile **bak** (.) mówi **vo** osiemdziesiąt litry czy siedemdziesiąt pięć mówi (.) **pod naklonom** (.) tam taki **mużik** jeden mówi objaśnicie wy dla mnie durnego mówi co to znaczy **pod naklonom** ((laughter))
- S: ((laughter)) wiesz jak jeszcze mogą robić? jak te **krepeži** jest które trzymają **bak**—
- T: mhm
- S: — ich oba zrazu tak **vo nu— nu** nie na około tak weźmą (.) a **prjamo do ramy** (.) i z boku jego wydują (.) ni widać **nifiga**
- T: a trzyma się?
- S: jeżeli dobrze postawisz to będzie trzymać się
- T: to jaka waga bierz
- S: **nu** (.) a wyduwają to wiesz jak fajnie— fajnie wyduwają (.) jego zdejmują (.) bok nagrzewają (.) i **kompessorom** powietrze dają
- T: **nu** ja wiem to że **kompessorom** to (.) tak i ty u siebie w maszynie możesz **bak vydut'**
- S: **da** (.) **plastassovj bak**
- T: mhm?
- S: **plastassovj bak** mogą wyduć
- T: a czemu nie?
- S: później jak wydują
- T: nie (.) (tak) wyduwają (.) zwiększają baki
- S: byłoby mnie co do niego ło— zalewać tylko
- T: tylko że (.) w tych żesz Pasatach po sto—
- S: w **v del'finax**?
- T: — w:: starych (.) w trzecich
- S: **nu da** (.) **del'fin**
- T: mhm (.) **del'fin** B czwarty (.) liczy się
- S: mnie zdawało się że B trzeci

- T: mhm ((negating))
- S: **nu** i co tam robio?
- T: tam biero na przykład też oni wyduwajo (.) s— sto litrów (.) sto dwadzieścia robio
- S: **pożalujsta** (0.2) **da** ()
- T: co ty chcesz?
- S: **nastojaščije mašiny** trzeba popatrzeć
- T: BMW?
- S: **da**
- T: BMW to nie maszyna
- S: a co to takiego?
- T: mhh (.) powiem tobie później
- S: ((laughter)) oj Boże/ **oj Boże** (.) **kul’turnyj naš—našelsja**
- T: ja tobie pokaža **nastojašč**-a maszyna
- S: o:: **da bros’** (.) ona koło Opery stoi
- T: nie (.) tam nie maszyna jak dla mnie (.) jaka mówisz?
- S: DAS auto
- T: DAS auto
- S: a Niemcy normalnie przymyśleli (.) dasz auto? ((yawning)) (0.8) ciekawie dzisiaj zdam czy ni zdam **teorija**
- T: a tam **vrode by** trzydzieści pytań benci
- S: **da** ((dismissive)) (.) **vo** (.) **bljaxa** (.) Miron (.) **odolži a?**
- T: dobrze (.) **bez problem** (.) co to jest pięćdziesiąt pięć tysięcy (.) noc nie pospać
- S: () (tarifie)
- T: **nu** wiesz jak na takiej maszynie (.) na nowej (.) to i ja jeździłby
- S: **blin** (.) cztery V.D
- T: ale ty— ale tylko nie trójka
- S: **bljaxa** cztery V.D (.) ()
- T: ale tylko nie trójka

- S: czemu? trójki nie podobajo się?
- T: mh-h ((negating))
- S: czemu?
- T: aj ni wiem (ja) (.) małe
- S: baba jezdziwszy na maszynie tej
- T: to ja mówia (.) dla bab— dla baby taka— dla dziewczyny
- S: **bljaxa** (.) ty dla swojej dzieczyny taka maszyna kupiłby?
- T: **да**
- S: t— **bljaxa** (.) **lituxa** ja tylko zdjąłby (.) nic
- T: a co to (.) **lituxa** tobie
- S: ładne koła
- T: mhh?
- S: koła ładne i **vse** (.) wczoraj () dostał
- T: mhh
- S: prz— trzy dostał normalnie (.) z trzeciej cuzć nie zwałił się
- T: osiem— osiem dwieście za **devjanosto devjatyj god** (.) oddać
- S: to po twojemu drogo?
- T: **konečno** (.) bardzo drogo
- S: (0.4) () **zabył voobše** (.) czekaj (.) gdzie to
- T: nie wykluczaj
- S: kto powiedział że wykluczać bendam (.) **voobše**
- T: to możesz za takie piniędzy wziąć—
- S: dobrego **gol'fa**
- T: **da** (.) na trzy d— nie (.) **daže Audi** (.) czekaj (.) ile euro wychodzi osimem dwieście?
- S: gdzie:: dwa:: trzy:::sta pięćdziesiąt
- T: **nu to vot** (.) to **gde-to** za takie piniędzy możesz wziąć sobie **naprimer** A trzecia (.) Audi
- S: nie (.) mnie A trzecia nie podoba się

- T: on a bendzi i— motor lepszy i **obsłużka** taniejsza po zdrównaniu **vo** na jej rok będzi lepszy dwóch cięczny jaki trzeci czwarty (0.4) albo **xot'** tego samego **Opelja** mozesz wziąć (.) jakiego Zefira czy co (.) i bendzisz jeździć (.) albo Szarana
- S: ja już lepiej na te pieniądze sobie nakupiam zefirów
- T: aha (.) jeszcze **dokinut'** jaki **naprimer tam/** tam tysięcy dwa(.) Szarana mozesz wziąć z jakim dobrym **naprimer** już nowym— bendzi nowa **morda** (.) **vse** (.) **televizor** już bendzi nowy już **tipa— vo** (.) musi **rešetka** ładna
- S: aj (.) sama—
- T: be— motor jaki **naprimer** bendzi już **odin devat'** (.) **vosem'desjat pjat' kilovatt** jeszcze **dokineš'** jakich na przykład pięć tysięcy to bendzi **devjanosto šest' kilovatt**
- S: Miron (.) **ne soblajnaja** (.) **bo— ja vse ravno Opelja bol'se nikogda ne kuplju**
- T: ja **vot** ni wiem (.) tak żeby co tylko ni Golf
- S: czemu tak?
- T: już mnie— bo rozumiesz (.) tak nadojadszy
- S: () ()
- T: do garażu ida płacząc (.) ni moga patrzeć na swoja maszyna
- S: ja myślał że ty/ **ty svoego Gol'fa ljubiš'**
- T: lubia lubia ale kiedy on mnie lubić zacznij
- S: **nu—**
- T: bendzi 'bam'
- S: **kogda nibud'** (.) **kogda baby načnut ljubit'** (.) **togda i Golf poljubit**
- T: **baby/baby** mine liubio (.) ja ich—
- S: **da** (.) **da** (.) **ty passkazyval** (.) cały Golf **v bolote**
- T: cicho
- S: sześćset **evro za kuzov** (.) bez— bez drzwiów ((clicking through websites)) **blin**(.) **angličanka**
- T: tak (.) zaraz (.) napisza sms-ka
- S: pisz (.) nie spieszajsia (.) jeszcze **vagon vremeni** (.) **blin**

- T: ((whispering)) ja chcem pi-pi
- S: idź
- T: zara (.) cyrk (.) ciekawie dzisiaj teoria— **nu** musza zda'c (.) ale **esli** ja zdam to nie zdziwcie się że **liven'** bendzi
- S: już **liven'** (.) bez ciebie
- T: bo co B1 zdawał **nazad** wracalisia był **liven'** (.) co **B kategorija** zdawał też **liven'** był
- S: a **xot'** zdał?
- T: **da** (.) **nu** ja **vse** z pierwszego razu zdawał to mnie dzisiaj taka odpowiedzialność że ja musza znowu **vse** z pierwszego razu
- S: kto tobie to powiedział?
- T: ja sam sobie powiedział
- S: a w lusterko patrzył? z— zdasz z pierwszego razu (.) **ne vozmuščajsja**
- T: nie **nu prosto** jak (.) tam zdawał z pierwszego razu (.) i tu musza (.) **prosto** mnie ciekawie (.) (()) żeby **klass byl/był**
- S: u ciebie **glaz uže est' podbityj**
- T: **klass a ne glaz** ((laughter)) dam w ucho!
- S: j— jeżeli zdasz to **v ponedel'nik na fure prijedeš'**
- T: **s teorii** (.) **da?**
- S: takie i sprawy
- T: mhm (0.12) zmęczył się ja już
- S: nie ty jeden (.) trzeba iść **pospat'**
- T: **pospat'** (.) ja wczoraj z garażu musi wpół do dwanastej wylazł tylko
- S: póki gotował swoja **kulisa?**
- T: nie tak póki— tak żeby **kulisa** zdjąć **glušitel'** trzeba było w dwóch miejscach odkroić (.) trąba wybrać (.) **zaščita** zdjąć (.) i wtedy tylko **kulisa**—
- S: odkroić?
- T: **da** (.) bo był dla jednego zdawszy jak jeszcze dobre czasy u mnie byli (.) kiedy pieniążków miał (.) to był zdawszy (.) on mnie i od **katalizatora kreplenije** wykroił **prosto** trąba (.) i wdrugim miejscu gdzie musiało być **xomut** też wykroił **prosto** trąba **vvaril** (.) jeszcze trąba cieńsza **vvaril** i ona **postojanno** mnie **lomala** (.) widzisz jak dobrze

- S: bardzo
- T: ehh (.) ja wziął ta trąba **perevaril** (.) druga postawił na— wiesz (.) nie domyślał się że **kulis** u mnie kiedy to może połamać się (.) **kulisa** połamała się (.) co robić (.) z jednej strony () (.) z drugiej strony trąba zdjął (.) **zaščita** zdjął (.) tak i **kulisa** zdjął (.) pojechał do jednego magazynu (.) do Auto Aibé (.) tam kupił jeden ten **xomut** (.) tam kupił ja też dobra pięćdziesiątka **tipa** (.) poszedł do ‘Interkarsu’ mówia mnie taka sama dajcie (.) on dla mnie dał ja i nie popatrzył (.) tam pięćdziesiąt cztery było (.) mnie trzeba było **vnutennij** pięćdziesiąt **diametr** (.) on dał dla mnie sześćdziesiąt (.) stawić ta trąba zaciskiwać za mały na () tego (.) co robić (.) wziął **bulgarko** skroił skroił skroił tam te żelazki (.) **nu** tam formy jeszcze ido takie że ładnie **tipa** zakłada się (.) **vse** żeb równo (.) to wziół w jednej **uglubil** (.) z drugiej troche też skroił wziół **nu** (.) żeby mniej tego (.) ta skroił patrza **ljuks**
- S: **normal’no** (.) **molodec**
- T: **da** tak i—
- S: ty ni wiesz co powiedzieć (.) ja ni wiem— ni wiem jak ciebie dalej słuchać (0.2) **davaj Saniok, ty možeš’, ty možeš’** (.) **ne katis’ nazad** (.) **dvigaj v pered** ((laughter))
- T: ((laughter)) () **vse**
- S: **zachodi v svoju** (.) popatrzaj co tam robisie ((about e-mail))
- T: zara przewierzym (.) **devuški** muszo pisać mnie (.) **netu** (.) ja siedział w tej (.) na [LT] **pirmos pagalbos** ta siedziała (.) z boku przyleciała (.) myślał **San’ka** wezma z boku posadza (.) ten **Saniok** póki tam/ **tam gde to** na wronów patrzył patrza **devka** z tyłu leci (.) mówia dobrze (.) chodź (.) to **Saniok** siedział koło **manekena** (.) i **tetka** tam jemu **vs**— ona pociska (.) on pociska (.) i dobrze
- S: **kto pro tebjja vspomnil** (.) **čto rešil napisat’**
- T: **aj tut** musi **vse nasčet** tego [LT] **baliukas**-a jeszcze
- S: a ty jeszcze nie był? myślał już był u ciebie
- T: nie (.) jutro
- S: jutro?
- T: **da**
- S: u::: (.) **p’jan’ ty/ty** (.) chcesz **Opelja odolžu?** **pojedeš’** (**pustiš’**) **tam**

- T: **da:** (.) dziękuje (.) rozwali się po tej drodze (.) a tam fajne miejsce jest (.) zajeżdżasz— **tipa** na ta wioska zajeżdżać (.) a po lewej— a po lewej stronie jest eh:: czekaj (.) dobrze (.) tam **tipa** zajeżdżasz po lewej stronie taka **ploščadočka** (.) na ta **ploščadka** zajeżdżasz i (.) jeziora widać **vse** ładnie (.) i zamek
- S: w Trakai **da**?
- T: **da**
- S: droga dobra? można dojechać?
- T: tam droga dobra tylko na ta wioska już jak skręcać to tam
- S: Wilejka?
- T: kolega piszy (.) z Kaunasa jedzi (.) on ż na/ **na policejskiego** uczy się
- S: mhh
- T: u nas w czternastym czy pi— (.) nie (.) **vrode by** że w piętnastym roku bendo mieniać uniformy/ **uniformy u policejskich** (.) tego **tipa** bendo sinie (.) jak u wszystkich (.) jak **v Evrope**
- S: uhh **evrostandart**
- T: bendo tam wiesz (.) **pokručę** robić (.) tam [LT] **su Liet.** (.) tam [LT] **apšvietimą** (.) **vse takoe**
- S: żeby **menta s kilometra** mog by **zametit**?
- T: **da** (.) **po ljubomu u nas tut** połowa potem bendzi— ni bendzi widzieć (.) przyjado **kolchozniki** jakie takie **vo** (.) **kolchozniki** powiedzo kto ty taki?
- S: ty te— **ja ment** (.) **nifiga**
- T: ja jak zawsze za granica do Polski czy gdzie jedziesz i zrazu sobie tylko do głowy wbiwasz (.) **policejskije** maszyny inaksze (.) i oni inaczej wyglądajo
- S: **da** (.) pamiętam **v Italii menty koroče vo v takich vo sapogach chodjat do sjuda**
- T: jak u mnie jest te **voenskije** też **tipa** takie **kerzy** (.) to u nich musi też w tym coś **vrode** (.) **tipa kerzy**
- S: jak babskie **tufli**
- T: babskie **tufli**?
- S: jak babskie **tufli** tylko **bez kablyka**

T: **nu** musi **ponjal**

S: **u nich do sjuda**

T: ichne **uniformy** nieładne