

Language Contact: Morphosyntactic Analysis of Surzhyk Spoken in Central Ukraine

Kateryna Kent
University of Minnesota – Twin Cities

1 Introduction

The term ‘surzhyk’ comes from Russian denoting a mixture of wheat and rye flour; adulterated wheat flour (Ushakov 1940). Now Surzhyk is used to denote the Russian-Ukrainian language mixture. The word has a negative connotation and opposes the language mixture with varieties of language perceived as pure, such as the standard variety. According to Bilaniuk (2004), current public discourse associates Surzhyk with parochialism, lack of education, and a low culture.

The purpose of this research is to categorize this spoken linguistic variety by examining its morphosyntactic structure, and to identify what elements in the data are indicative of language mixing and what can be attributed to other contact phenomena such as codeswitching and lexical borrowing. Although in public discourse the term Surzhyk is used broadly to indicate anything from occasional borrowing to language mixing, in this paper I will give a more precise characterization of Surzhyk by evaluating its structure with respect to two language classification models: Matrix Language Frame model (Myers-Scotton 2002) and Auer’s continuum (1999).

Since there exist some other Ukrainian dialects not generated by contact with Russian, ethnographers and anthropologists use the term Surzhyk to indicate a language variety in which the grammar of Ukrainian — phonology, morphology, syntax, and lexicon — contains Russian-influenced elements which are not present in the Standard Ukrainian language (Flier 1998, Bilaniuk 2004). Long coexistence of Ukrainian and Russian and language policies implemented by the Russian Empire and the Soviet government created favorable conditions for development of the mixed language variety.

In the second section of the paper I will identify the differences between Russian and Ukrainian that are relevant for the analysis of Surzhyk. In the third part I will highlight previous research done on Surzhyk. In the fourth part I will describe two theoretical models which will be used for the analysis of Surzhyk. In the fifth part of the paper I will introduce the methodology of the research (5.1); detail the analysis of Surzhyk collected in Central Ukraine (5.2) by examining agreement markers (5.2.1), verbal derivational affixes (5.2.2), discourse markers (5.2.3), verbs of motion (5.2.4); and talk about codeswitching versus lexical

borrowing (5.3). In the sixth part I will discuss sociolinguistic factors to be taken into account when analyzing a mixed language variety. Finally, in the seventh part, I will present a conclusion.

2 Differences between the Russian and Ukrainian languages

One of the problems in classification and analysis of Surzhyk lies in the fact that two closely related languages, Russian and Ukrainian, comprise its core. For instance, word order, case systems, verb conjugation patterns, etc. are very similar. It might be difficult to determine what the source language of some morphemes in a mixed utterance is.

Both Russian and Ukrainian originated from the Proto Slavic language. The 6-7th centuries are the approximate period of formation of Old Russian and Old Ukrainian as two distinct languages (Ohienko 2001, Shevelov 2002). Although both languages retained a number of similarities - due to geographical proximity, common ancestry, and prolonged language contact - there are significant differences to be pointed out.

According to Tyshchenko (2000), the lexicons of Russian and Ukrainian differ by 38%. To compare with other European languages belonging to a single language family, Spanish and Portuguese differ by 25% and Spanish and Italian by 33% (Tyshchenko 2000:266–267). The remaining 62% of the Ukrainian lexicon contains 44% of lexical units morphemically identical to Russian, and 18% morphemically similar to Russian (Bilaniuk & Melnyk 2008). An example of two different lexical items is a word for flower in (1):

- (1) a. /kvitka/ (Ukrainian)
 ‘flower’
 b. /tsvətok/ (Russian)

The stem of the verb *to sleep* /spa/ is identical for both Ukrainian and Russian, while the word for *guest* (2) differs only in a stem vowel which makes it morphemically similar in the Nominative case and identical in other cases. I will not consider /fi/ and /g/ to be two separate phonemes because they are derived from the same Proto Slavic form (Shevelov 1977).¹

- (2) a. /fiistⁱ/ (U)
 ‘guest’
 b. /gostⁱ/ (R)

Both Ukrainian and Russian have a number of stems that differ only in one phoneme as in ‘old man’:

¹For the purposes of comparing Russian and Ukrainian, the two phonemes are diachronic representation of the same Proto Slavic form.

- (3) a. /did/ (U)
 b. /ded/ (R)

The reason for differentiation in (2) and (3) is historical sound change of etymologically related morphemes in both languages (Pugh 2007). The phonological systems of Russian and Ukrainian are different in both phoneme inventory and phonetic detail. The Ukrainian vowel system has six vowel phonemes, while Russian is a five-vowel system, as shown in Figures 1 and 2 (Bilous 2005, Zilynskyj 1979).

Figure 1: Phonemic Inventory of Ukrainian

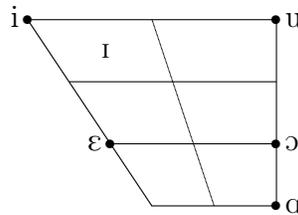
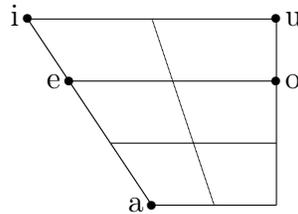


Figure 2: Phonemic Inventory of Russian



Unlike Russian, Ukrainian does not have a voiced velar stop /g/ in Ukrainian words: it appears only in loanwords and assimilated loanwords as in (4):

- (4) a. /grmɕɔɫɪ/ (Ruthenian loanword into Standard U)
 ‘sleigh’
 b. /gudzɪk/ (Polish loanword into Standard U)
 ‘button’

Instead, the glottal fricative /ɦ/ is used in native Ukrainian words (Shevelov 1977). In other words, in common east Slavic words Russian /g/ systematically corresponds to Ukrainian /ɦ/, as shown in (5):

- (5) a. /ɦɪɾɑ/ (U)
 ‘game’

- b. /igra/ (R)
 ‘game’

In Russian the vowel /o/ is reduced to [a] or [ə] in an unstressed syllable. This process is known as *akanie*. In Ukrainian this reduction of /ɔ/ does not occur. Instead, /ɔ/ surfaces as [o] in unstressed syllables. Example (6) demonstrates pronunciation of the Russian and Ukrainian words for *cow* although orthographically these words are identical in Russian and Ukrainian (Bilous 2005, Bilaniuk & Melnyk 2008, Jones & Ward 1969).

- (6) a. [kor'ɔvɑ] (U)
 /kɔrɔvɑ/
 b. [kar'ova] (R)
 /korova/

Another phonological process that differentiates Russian and Ukrainian is final consonant devoicing (Jones & Ward 1969). Russian voiced consonants devoice in word-final position, while in Ukrainian final devoicing does not occur as in the word for ‘old man’ in (7).

- (7) a. /did/ (U)
 b. /det/ (R)

Other phonological differences between Russian and Ukrainian include the palatalized fricatives /ʃ/ and /ʒ/ and the affricate /tʃ/ in Russian. In Ukrainian these phonemes are either hard (non-palatalized) or semipalatalized (Bilaniuk & Melnyk 2008).

In addition to differences in their phonological systems, Russian and Ukrainian also diverge in morphology and syntax when it comes to the case system, gender, noun declensions, tense, etc. Russian has six cases: Nominative, Genitive, Dative, Accusative, Instrumental, and Locative. Ukrainian shares these six cases with Russian, but it also retained one additional case from Proto Slavic: Vocative (Pugh & Press 1999, Ponomariv 2001). The Russian Vocative case was lost in the course of historical development. In Modern Russian Vocative case is retained only in the word for God (Bilaniuk & Melnyk 2008), for all other words the Nominative case is used, as illustrated in Table 1.

Table 1: Vocative Case in Russian and Ukrainian

Case	Russian		Ukrainian	
Nominative	/bog/ ‘God’	/zemlʲ a/ ‘earth’	/bɔfi/ ‘God’	zemlʲ a/ ‘earth’
Vocative	/bɔʒ-e/ ‘Oh, God’	/zemlʲ a/ ‘Oh, earth’	/bɔʒ-ɛ/ ‘Oh, God’	zeml-ɛ/ ‘Oh, earth’

The Ukrainian language can form the future tense in two ways: by means of a modal verb or by suffixation. The Russian future tense is formed only by means of a modal verb.

- (8) a. /budu pisatɪ/ or /pisatɪ-mu/ (U)
 be.1SG.FUT write.INF or write.INF-1SG.FUT
 ‘I will write.’
 b. /budu pisatʲ/ (R)
 be.1SG.FUT write.INF
 ‘I will write.’

The viewpoint aspect of the two languages is similar. It is conveyed mainly through verbal affixes that carry some additional meaning besides a mere perfectivizing function. There are 27 verbal prefixes in Russian and 24 in Ukrainian (Ozerova 2003). The number of verbal prefixes, the meaning they convey, and their compatibility with lexical verb classes differ in Russian and Ukrainian.

- (9) a. /po-jihatɪ/ (U)
 PREF-go.INF
 ‘to leave’
 b. /u-jehatʲ/ (R)
 PREF-go.INF
 ‘to leave’

In (9), the Russian and Ukrainian verbs convey the same spatial characteristics (moving off/away) by two different prefixes.

The case assigned by some verbs differs in Russian and Ukrainian. Thus, in Russian (shown in (10)) the verb *to thank* requires the Accusative case, while in Ukrainian (shown in (11)) *to thank* requires the Dative case.

- (10) a. /dʲakuvatɪ/ (U)
 thank.INF
 ‘to thank’
 b. /dʲakuvatɪ družʲ-am/
 PREF.thank.INF friends-DAT
 ‘to thank friends’
- (11) a. /blagadaritʲ/ (R)
 thank.INF
 ‘to thank’
 b. /blagadaritʲ druž-ej/
 PREF.thank.INF friends-GEN
 ‘to thank friends’

Another important distinction is the difference in case assignment by prepositions in Russian and Ukrainian. For instance, the preposition *po* ‘through’ assigns the Prepositional case to plural nouns in Ukrainian and the Dative case in Russian, as in (12):

- (12) a. /jihati po mist-**ah**/ (U)
 go.INF through city-PL.PREP
 ‘drive through cities’
- b. /jehatⁱ po garad-**am**/ (R)
 go.INF through city-PL.DAT
 ‘drive through cities’

Although genetic and typological proximity and centuries-long language contact brought similarities to Russian and Ukrainian, differences in morphology, syntax and phonology of these two languages are an important reference point for the analysis of such a language contact variety as Surzhyk.

3 Existing Research on Surzhyk

The origin of Surzhyk is connected with development of the Ukrainian language and Ukrainian-Russian language contacts caused by geographical proximity and colonization of Ukraine. Scholars agree that Surzhyk, as a linguistic phenomenon separate from Ukrainian and Russian, dates back to the 17th—18th centuries (Masenko 2008). Central, Eastern, and Southern provinces of Ukraine were annexed to the Russian Empire at that time. The Russification policy of the Russian Empire hampered development of the standard Modern Ukrainian language and created favorable conditions for creation of Russian-Ukrainian mixed language variety (Masenko 2008, Flier 2008, Stavytska & Trub 2007).

Although Surzhyk has been around for several centuries, it became the object of linguistic, anthropological, sociological and sociolinguistic research only in the last two decades, after the proclamation of Ukrainian independence in 1991. Since this linguistic phenomenon has not been extensively researched, there is no agreed upon classification of what constitutes Surzhyk or consensus on whether it is a single variety or a system of contact phenomena ranging from lexical borrowing to code-switching to language mixing.

The main disagreement among scholars is the nature and number of Surzhyk(s). Thus, some (Flier 1998, Stavytska & Trub 2007; Vakhtin, Zhironkina, Liskovets & Romanova 2003) agree that there is only one linguistic variety called Surzhyk. However, they vary on what this variety is. Others (Bilaniuk 2004) propose several different types of Surzhyk. There is yet a third group of researchers (Serbenska 1994, Masenko 2008) who maintain that Surzhyk is not a separate language but simply an ad-hoc combination of Russian and Ukrainian morphemes without any consistent structure.

Bilaniuk (2004:415) proposed a typology containing five subcategories of Surzhyk: urbanized peasant Surzhyk, village dialect-Surzhyk, Sovietized-Ukrainian Surzhyk, urban bilinguals’ Surzhyk and post-independence Surzhyk. Bilaniuk’s typology is helpful in that it

allows us to pinpoint stages of Surzhyk development. The table (2) below illustrates the five Surzhyk prototypes.

Table 2: Five Surzhyk Prototypes (Bilaniuk 2004:415)

Types of Surzhyk	Specific Description	Rural–Urban Context	Era	Direction of Influence
Urbanized peasant	Working class urbanized U peasants	rural to urban	19th c. to present	R onto U base
Village dialect	U villagers in contact with R administration and media	rural	19th c. to present	primarily R onto U base
Sovietized Ukrainian	Codified U with planned R influence	urban (institutional)	1930s to present	R onto U base
Urban bilinguals'	Urban bilinguals with either native language	urban	Soviet and Post-Soviet	both directions
Post-Independence	Russophone urbanites newly using Ukrainian in public	urban	Post-Soviet	both directions

Flier (1998, 2008) disagrees with Bilaniuk, stating that there is only one linguistic variety termed Surzhyk. However, the author allows two distinct subcodes within a single structure. Flier concludes that "... there are two different systems or codes of Surzhyk, namely, Russian-Ukrainian Surzhyk, a Russian base with a Ukrainian admixture, symbolically R/U; and Ukrainian-Russian Surzhyk, Ukrainian base with a Russian admixture, symbolically, U/R ... the Russian admixture in the latter is itself Ukrainianized; therefore, Ukrainian-Russian Surzhyk is more accurately symbolized U/R^U" (Flier 2008:44). The author argues that Surzhyk demonstrates a set of structural features, such as lexical extensions, lexical calques, lexical transfers, syntactic calques, morphological selection, morphophonemic attraction and phonological integration, that allows him to classify it as a monolithic phenomenon (Flier 2008:45). This classification ascribes Surzhyk to mixed languages. However, it does not fully explain how two separate subcodes can play into one monolithic mixed language. Although the author's arguments against multiple prototypes of Surzhyk proposed by Bilaniuk are very sound, he does not use any spoken data to support his claim. The literary data used in the argument was created by writers of various epochs who were educated and proficient either in Ukrainian or both Ukrainian and Russian. They used Surzhyk for stylistic effect in their works. The analysis of spoken data presented in this paper can prove very useful to test the hypothesis proposed by Flier.

Vakhtin et al. (2003) showed that Surzhyk is a Ukrainian-based variety containing Russian lexical items and mostly Ukrainian morphosyntax. The researchers did not take any particular stance on the classification of Surzhyk. However, they did a meticulous analysis

of the language of origin for phonological, morphological, syntactic and lexical features of Surzhyk. For example, the data shows that the verbal prefix *pri-* is supplied by Ukrainian and is attached to mostly Russian stems, as in the lexeme /prijezatʲ/ ‘to arrive’. Although this study was the first one to use actual spoken data for the analysis, the analysis itself was mainly descriptive. The authors did not propose any general conclusion about the structure and constraints of Surzhyk. However, the list of morphemes and their origin provided in the analysis is very useful for comparative study of spoken data in different regions.

Stavytska & Trub (2007) define Surzhyk as a colloquial variety that originated as a result of Russian-Ukrainian bilingualism with diglossic relationships. Surzhyk is a result of systemic interference of phonetic, morphosyntactic and lexical levels of the grammatical system. Surzhyk is represented through Surzhyk lexemes that are incorporated into Ukrainian or Russian grammars. The authors identify interference at different levels of grammatical structure:

- Lexical and phonetic interference: Russian lexemes introduced into Ukrainian are pronounced according to Ukrainian phonological rules (relevant segments are bolded), ex., [b**olʲ** n**itsʲ**a] ‘hospital’ instead of [likarnʲa], a Ukrainian equivalent or [balʲnitsa], a Russian equivalent; [p**olezno**] ‘usefully’ instead of [korisno] (U) or [palezna] (R); [pozav**trikav**] ‘he had breakfast’ instead of [posnidav] (U) or [pazavtrakal] (R) (Stavytska & Trub 2007:74).
- Lexical and grammatical interference: Russian lexemes appear with Ukrainian agreement markers (bolded segments): [ponʲ**av**] ‘he understood’ instead of a Russian past tense agreement marker *-l* [ponʲal]; [poluch**aje**] ‘he/she receives’ instead of a Russian present tense agreement marker *-jet* [paluchajet] (Stavytska & Trub 2007: 75).

Although researchers do not consider Surzhyk to be codeswitching, they state that the following codeswitching scenarios are possible: Russian-Surzhyk, Russian-Ukrainian-Surzhyk, or Ukrainian-Surzhyk.

Masenko (2008), following Serbenska (1994), argues that Surzhyk is not a separate linguistic variety but simply a chaotic mixture of Ukrainian-Russian idiolects. She based this conclusion on her descriptive analysis of spoken data as well as on some recorded archived data from the beginning of the 20th century. The author maintains that dominance of Russian elements on phonological, morphosyntactic and lexical levels allows her to classify this linguistic variety as Ukrainianized Russian rather than Surzhyk (Masenko 2008).

The lack of consensus in the present scientific discourse on what can be classified as Surzhyk can be eliminated through a systematic analysis of the Surzhyk corpus. So far only a few researchers (Vakhtin et al. 2003, Bilaniuk 2004) collected and analyzed spoken data and demographic information of research participants that allowed them to assess surface features of this contact phenomenon. Others (Masenko 2008, Serbenska 1994) relied on sociopolitical and historical discourse to reach their conclusions about the origin and classification of Surzhyk. Yet another group (Flier 1998, 2008) analyzed literary works by the authors in

which Surzhyk is used as a stylistic device.² The drawback of such a method is the artificial nature of the language. Most likely, the authors of literary works are Russian-Ukrainian balanced bilinguals who are not native Surzhyk speakers and might not convey true features of this variety.

4 Frameworks for analysis of Surzhyk

There are several approaches to analyzing a mixed language variety. I will outline two of the approaches that will be used as theoretical tools for the analysis of Surzhyk in this paper: Matrix Language Frame model (Myers-Scotton 2002, 2003) and Auer's continuum (Auer 1999).

The Matrix Language Frame model (MLF) proposed by Myers-Scotton (2002, 2003) explains the contact phenomenon through the interaction of competence and performance by treating the lexicon and morphosyntax of a language within a single perspective. According to the MLF, only one language is the source of the abstract morphosyntactic frame in a bilingual clause (the Matrix Language) and the other participating language (the Embedded Language) must agree with structural requirements stipulated by the Matrix Language (Myers-Scotton 2002, 2003). Example (13) illustrates a bilingual IP from Surzhyk³⁴:

- (13) vse na *aharod-i* bulo
 everything in kitchen.garden-PREP be.NEUT.PAST
 'everything was in the kitchen-garden'

The IP in (13) shows that the main language supplying the abstract morphosyntactic frame is Ukrainian: the pronoun *vse* and the verb are supplied by Ukrainian. Although the noun in the PP 'in the kitchen-garden' comes from Russian, it has a Ukrainian Prepositional morpheme *-i*, required by the verb *bulo*, thus observing the well-formedness principle for the morphosyntactic frame. If this PP had been structured according to the rules of Russian, the morpheme *-e* would have been used, as in (14). But this form is unattested in Surzhyk.

- (14) na *agarod-e*
 in kitchen.garden-PREP
 'in the kitchen-garden'

The MLF provides mechanisms to distinguish language mixing from codeswitching by differentiating between early system morphemes (derivational affixes, plural markers) and late system morphemes (agreement markers). Discourse markers, as well as thematic roles

²For instance, one of the analyzed works was written at the end of the 19th century by a famous Ukrainian writer Starytsky who was highly educated and wrote a number of works in Ukrainian.

³I am using the Library of Congress Romanization style to illustrate the examples.

⁴Russian or Russified elements are italicized; Ukrainian elements are not marked; the elements under discussion are bolded.

assigners and receivers, are examples of content morphemes. Early system morphemes supply conceptual information about the heads occurring in the same immediate maximal projection. Late system morphemes look for the information about their form beyond their maximal projection, such as subject-verb agreement, case markers, etc.

The MLF uses its classification of different morpheme types to define three language contact outcomes: classic codeswitching, composite codeswitching and language mixing. In classic codeswitching the morphosyntactic frame and late system morphemes are supplied by one language, the Matrix Language, and only content morphemes come from the Embedded Language. In composite codeswitching the morphosyntactic frame and some late system morphemes are supplied only by the Matrix Language, but content morphemes and some late system morphemes come from the Embedded Language. In a mixed language, late system morphemes are supplied by the Embedded Language in at least one constituent type or there is a loss of the late system morphemes in the Matrix Language in at least one constituent type (Myers-Scotton 2002).

Auer (1999) proposed a typology to differentiate between types of language interaction such as code switching, language mixing and fused lects. All of these phenomena should be viewed as points on a continuum with codeswitching and fused lects being the extreme poles and language mixing being the halfway point between them.

Code Switching	Language Mixing	Fused Lects
●		●

For codeswitching, a speaker is usually bilingual in both languages involved in the codeswitching. He is aware of the differences between two (or more) codes. According to Myers-Scotton (2002, 2006), codeswitching is usually volitional and is triggered by a situation (discourse-related switching) or some characteristics of an interlocutor (participant-related switching). In language mixing it is impossible to determine the language of communication due to frequency of intrasentential switches (Auer 1999). A speaker does not have a preference for using one language at a time. In fused lects, language mixing is obligatory. Surzhyk is said to include phenomena at various points of this continuum (Bilaniuk 2004, 2005). A more precise characterization of this variety will allow me to verify or narrow current definitions.

5 Analysis of Surzhyk

As I have outlined above, some questions about Surzhyk that are debated or have not yet been researched still need an answer. In this paper I will answer the question whether Surzhyk is a case of codeswitching, language mixing or lexical borrowing. To answer this question, a systematic research program on spoken Surzhyk data is needed. I maintain that the scholarship on Surzhyk will greatly benefit from a structural approach to this contact

phenomenon, which will allow us not only to describe the surface features but also conclude why underlying morphosyntactic structures of two languages interact with each other the way they do.

However, looking purely at the structural features of the language would be misleading. In this paper I will also consider sociolinguistic factors; that is, language of preference of my research participants and their ability to codeswitch.

5.1 Methodology

The linguistic demographics of Ukraine vary significantly in different regions. This is attributed to different historical factors which influenced development of Ukrainian identity, and different literary and language traditions. Researchers divide Ukraine into four parts according to its cultural and linguistic features: East, West, South and Central Ukraine. The largest number of Surzhyk speakers was recorded in eastern, southern and central parts of the country (Vakhtin et al. 2003). There is no research confirming Ukrainian-Russian Surzhyk in Western Ukraine. However, the public discourse points out to the number of lexical borrowings from Polish (Flier 2008).

I conducted field research in summer of 2009 in the Kiev region, which is in Central Ukraine. Nineteen (19) participants agreed to take part in the research. The participants were both from urban centers and rural areas. Before and during recording, I addressed participants either in Russian or Ukrainian. I switched languages during the conversation to decrease participant's accommodation to my language in their answers and also to eliminate any ideologically charged conclusion about which language is more important, Ukrainian or Russian (Myers-Scotton 2006, Bilaniuk 2004, Bilaniuk & Melnyk 2008). For instance, if I talk about Russian-Ukrainian bilingualism but I ask questions only in Ukrainian, I might be seen as considering Ukrainian to be more important and thus participants might try to accommodate their speech to match that assumption.

Notably, all of the participants reported speaking Surzhyk⁵ at home. None of the participants demonstrated a command of both Ukrainian and Russian. As the language of the interviewer changed from Ukrainian to Russian or vice versa, the participants maintained the same code throughout the interview.

5.2 Analysis of data and findings of the research

As discussed above, discourse analyses of literary works done on Surzhyk to date fail to capture larger generalizations or show underpinnings of language contact outcomes. Thus, I adopted the MLF model introduced by Myers-Scotton (2002, 2003) and described above as a convenience tool for my analysis. There are several advantages for using this model when analyzing contact languages like Surzhyk: (1) analysis within this framework will allow me

⁵Some participants call their language 'mixed Ukrainian' or 'this type of Ukrainian' instead of Surzhyk. Although they may not call the variety they speak Surzhyk, they clearly differentiate it from both Ukrainian and Russian.

not only to classify this variety but also distinguish among classic codeswitching, composite codeswitching or language mixing, the types of contact phenomena characterized by the model; (2) since the MLF model operates by distinguishing different types of morphemes (content morphemes, early system morphemes and late system morphemes) it will be possible to adjust this instrument, if necessary, to account for closely-related languages, such as Russian and Ukrainian. Although this model is used only as a testing tool, as can be seen below, it does not work for this set of data because it forces the conclusions that do not make sociolinguistic sense.

Analysis of the spoken data of 19 research participants showed that the Matrix Language is Ukrainian and the Embedded Language is Russian since Russian supplies some of the content morphemes, while Ukrainian supplies morphosyntax in a mixed IP. Consider the following examples:

- (15) pid nas nihto ne podstra-**iuva**-v-sia
 under we.GEN nobody NEG adjust-2IMP-3SG.MASC.PAST-REFL
 ‘nobody accommodated to us [nobody adjust their schedule to accommodate us]’

In (15) the Russian verb for ‘*adjust*’ appears in an otherwise Ukrainian IP. Note that the verb has the Ukrainian secondary imperfective suffix *-iuva*. It also is inflected for the third person singular masculine past tense, which is conveyed by the Ukrainian morpheme *-v*. The reflexive morpheme *-sia* has the same phonological and morphological properties in both Ukrainian and Russian. So, in this case it is irrelevant what language it came from. But the fact that the secondary imperfective and agreement markers on the verb (these elements are late outsider system morphemes according to the MLF model) came from Ukrainian indicates that Ukrainian indeed supplies the morphosyntactic frame for the Matrix Language. Russian, then, is the Embedded Language.

Consider another example which illustrates similar behavior of late system morphemes in a different constituent type:

- (16) a. salo nikoly lysh-**n’oi** sol-i ne viz’-me
 smoked.fat never extra-GEN salt-GEN NEG take-3SG.FUT
 ‘Smoked meat will never absorb extra salt.’
 b. lishn’-e^ǐ sol-i (R)
 extra-GEN salt-GEN
 c. zaiv-o^ǐ sol-i (U)
 extra-GEN salt-GEN

In example (16a), another mixed constituent is demonstrated. In the NP ‘extra salt’ the adjectival stem is supplied by Russian, but the adjective-noun agreement morpheme for the feminine genitive ending comes from Ukrainian. This morpheme also belongs to the class of late outsider system morphemes according to the MLF. In (16b) and (c) the

Russian and Ukrainian NPs are demonstrated. Note that the stem vowel in the Surzhyk NP (16a) differs from the Russian stem vowel in (16b) because Russian content morphemes introduced to the Ukrainian morphosyntax are also subjects to Ukrainian phonology and are often pronounced in accordance with Ukrainian phonological rules. Although both languages have non-palatalized /l/ and palatalized /lʲ/ phonemes (see Table 3), the non-palatalized phoneme in Ukrainian usually surfaces in cognate words more frequently than palatalized as illustrated below:

Table 3: Distribution of /l/ and /lʲ/ in Russian and Ukrainian

	Russian	Ukrainian	Gloss
1	/lɪsɪj/	/lɪsɪj/	'bold'
2	/lʲɪpkʲɪj/	/lɪpkɪj/	'sticky'
3	/lʲɛpʲtʲ/	/lʲɪpɪtɪ/	'mould'

To recapitulate, the data analysis shows that the morphosyntactic structure of Surzhyk is Ukrainian with some content morphemes supplied by Russian and often subjected to Ukrainian phonology. The MLF analysis of Surzhyk data recorded in Kiev indicates that it is a case of composite codeswitching.

There are several pieces of evidence, which I will detail in the subsections below, that lead to the conclusion that Surzhyk is a case of composite codeswitching: most agreement markers are supplied by Ukrainian in a wide range of constituents (5.2.1); verbal derivational affixes that mark viewpoint aspect and add additional meaning to a verb are supplied by Russian on both Ukrainian and Russian verb stems (5.2.2); discourse markers are supplied by Russian, which suggests codeswitching at the discourse level (5.2.3) and some Ukrainian verbs of motion select for the Russian preposition (5.2.4).

While three out of four characteristics outlined above indicate that Surzhyk is codeswitching, I suggest that it is the last fact about government of some Ukrainian motion verbs that most critically indicates that Surzhyk is a case of composite, rather than non-composite, codeswitching rather than codeswitching. The relationships between a verb and the preposition it selects are established after all content morphemes with their semantic and pragmatic features as well as early system morphemes are assembled, thus making the preposition and its complement a late system morpheme. If all the motion verbs in mixed Surzhyk IPs had selected Russian prepositions, I could have argued that Surzhyk is indeed a mixed language. But since it is not the case, I have to settle for composite codeswitching.

5.2.1 Agreement markers

Recall our earlier discussion about the Matrix Language – Embedded Language opposition within the MLF model (Section 4) and its relevance to the definition of codeswitching. In the case of composite codeswitching most agreement markers (late system morphemes) are

supplied by the Matrix Language. In Surzhyk they are supplied by Ukrainian (the Matrix Language) in a wide range of constituents, as I have already illustrated in examples (13-16). Consider the examples below:

- (17) *sledu-che* selo
 next-NEUT village
 ‘next village’
- (18) *svoi slova taki interesn-i*
 own.PL word.PL such.PL interesting-PL
 ‘their own words so interesting’

In (17) the Russian participial stem ‘next’ takes a Ukrainian agreement marker. In (18) the NP ‘their own words so interesting’ features a content morpheme from Russian with Ukrainian agreement suffix.

5.2.2 Verbal derivational affixes

In cases of codeswitching, early system morphemes (derivational affixes) can be supplied by either the Matrix or the Embedded Language. Both Russian and Ukrainian have a set of verbal prefixes (derivational affixes) that mark viewpoint aspect and add some additional meaning about the progress of the action to a verb. In the Kiev data these prefixes are supplied by Russian on both Ukrainian and Russian verb stems. Examples (19)-(21) illustrate sentences with verbs that appear in Ukrainian IPs. All the verbs have Ukrainian agreement markers. However, verbal derivational prefixes are supplied by Russian. As I have mentioned earlier, Russian and Ukrainian share many common roots, which seems to be especially true in the class of verbs. But what makes this case interesting is the fact that even those Ukrainian verbs that do not share common roots with Russian still can appear with Russian derivational prefixes.

- (19) *my pry-vykl-y* tak
 we PREF-get.used-1PL so
 ‘We got used to this.’
- (20) *vona na-chyna-ie*
 she PREF-begin-3SG
 ‘She begins.’
- (21) *nimtsi u nas od-roblia-l-y*
 Germans by we.GEN PREF-work-3PL-PAST
 ‘Germans worked here.’

In sentences (19)–(20) verbal stems are shared by Russian and Ukrainian. In (21) the Ukrainian verb stem is used with Russian derivational prefix. Let us look at each example in more detail. In example (19) the stem *pryvykly* features the Russian prefix *pry-* and the stem *vyk* which is identical in both languages. The Ukrainian counterpart of this verb is *zvykly* ‘got used to’. Surzhyk speakers select the Russian verbal prefix to convey viewpoint aspect. A similar case is illustrated in (20) for the verb *nachynaje* ‘he/she/it begins’, which would have the verbal prefix *po-* in Ukrainian. In (21) the Ukrainian verb *odroblialy* is used with the Russian verbal prefix instead of its Ukrainian counterpart *vidroblialy*.

These data confirm that Surzhyk’s derivational verbal prefixes, which are classified by the MLF model as early system morphemes, come from Russian. This is consistent with the codeswitching phenomenon, as discussed by Myers-Scotton (2002).

5.2.3 Discourse markers

Surzhyk displays a number of mixed CPs, shown in (22)–(23), in which discourse markers are supplied by Russian.

- (22) *no v asnavnom po-ukraïns’ki balakaiut’*
but in general Ukrainian speak.3PL.PRES
‘But in general they speak Ukrainian.’

- (23) *patamu sho tse nepriestizhna bulo*
because it NEG.prestigious be.3NEUT.PAST
‘Because it wasn’t prestigious.’

The discourse markers appear as the C head of the CP. In some other languages, such as Chaupi Lengua (Quechua-Spanish mixture), similar surfacing of Spanish discourse markers triggers a change in morphosyntactic structure of the Matrix Language (Quechua) making the Matrix Language neither truly Spanish nor Quechua. This signifies the process of Matrix Language Turnover (whereby the former Embedded Language becomes the Matrix Language), allowing for classification of Chaupi Lengua as a mixed language (Myers-Scotton, 2003). But since Surzhyk is a mixture of two closely-related languages, discourse markers from either one of them do not have any significant effect on the morphosyntactic structure of Surzhyk overall and may simply indicate codeswitching at the discourse level.

5.2.4 Verbs of motion

I argue that preposition choice of the verbs of motion in Surzhyk is due to structural changes in the class of motion verbs induced by Russian grammar rather than codeswitching between Russian and Ukrainian. In Ukrainian, motion verbs indicating movement towards/into a place require the preposition *do* followed by an NP in the genitive case or the preposition *v* (*u* before consonants) followed by an NP in the accusative case. In Russian, the motion verbs require the preposition *v* with the accusative case. In Surzhyk the motion verbs are supplied

by Ukrainian but they appear only with the proposition *v (u)* followed by the accusative case. The preposition *do* is unattested in my data. Consider examples below:

- (24) *povezty v Rosi.iu*
 take.INF to Russia.ACC
 ‘to take to Russia’
- (25) *pishla v shkol.u*
 go.3SG.PAST to school.ACC
 ‘she went to school’

Since Surzhyk’s morphosyntactic frame is supplied by Ukrainian, one would expect these VPs to have PPs with the preposition *do* and the noun in the Genitive case as well as PPs with *v* and the Accusative case. But that is not the case in Surzhyk. So, examples (24)-(25) indicate that the part of morphosyntactic frame comes from Russian. Hypothetically, if the whole class of verbs or at least all of the motion verbs displayed a similar pattern, I could argue for Surzhyk to be a mixed language according to the definition of the MLF model. But this phenomenon is found only in cases of motion verbs that denote the movement towards/into a place.

5.3 Codeswitching versus lexical borrowing

Some scholars (Poplack 1980, MacSwan 1999) argue that any singly occurring lexical element with inflections from the source language is a borrowing. In this section I will demonstrate that the lexemes illustrated in this paper and containing a stem from one language (usually Russian) and affixes from another language (Ukrainian) are not Russian borrowings into Ukrainian because they are not used in the Standard Ukrainian language and they are not used by the Ukrainian-speaking population of Western Ukraine. There is, however, a number of Russian lexical borrowings into Ukrainian, some of which I will illustrate below,

- (26) *zavod*
 ‘factory’

The word *zavod* in (26) has been borrowed from Russian and functions alongside its Ukrainian equivalent *fabryka* ‘factory’.

In (27a) the borrowed deverbal noun has been morphologically and phonologically adopted to Ukrainian. Compare (27a) with its Russian counterpart in (27b),

- (27) a. *rozporiadzhennia* (U)
 ‘order, decree’
 b. *rasporiazheniie* (R)
 ‘order, degree’

The Russian prefix *ras-* was exchanged for the Ukrainian equivalent *roz-*. The Russian fricative \widehat{zh} became \widehat{dzh} in Ukrainian. The borrowing conformed to the group of Ukrainian neuter nouns of the second declension and acquired the Ukrainian neuter suffix *-ennia*.

I have demonstrated some lexical items in the examples throughout this paper that display similar behavior with Surzhyk lexemes, namely, that have Russian stem and Ukrainian affixes. Although Surzhyk lexemes conform to Ukrainian phonology most of the time, one should take into account lack of usage and acceptability of Surzhyk lexemes in the Ukrainian speaking part of Ukraine. Therefore, Surzhyk lexemes are not simple lexical borrowings from Russian into Ukrainian but rather the indication of the other language contact outcome.

6 Sociolinguistic factors and reservations about codeswitching

Up to this point I have only considered grammatical properties of Surzhyk and concluded that it can be characterized as composite codeswitching within the MLF model on the basis of the presence of late system morphemes supplied by Russian in some PPs. Following Thomason & Kaufman (1988), however, it is important to consider both the internal and external factors of language contact; that is, to analyze both grammatical structure and sociolinguistic characteristics. Consideration of sociolinguistic characteristics of the speakers poses a problem to the classification of Surzhyk as a case of codeswitching because its speakers did not demonstrate or report active command of both Ukrainian and Russian. The majority of the 19 recorded Surzhyk speakers reported speaking it at home as the main language of communication. Moreover, out of 8 recorded conversations involving a research participant and a Russian or Ukrainian speaker, the research participants did not switch between Surzhyk and Russian (or Ukrainian). This indicates that Surzhyk functions as a single code.

Another reason to classify Surzhyk as being closer to a mixed language is the existence of some forms distinct from both Russian and Ukrainian. In Surzhyk, numerals usually combine with nouns according to Ukrainian grammar. Consider the following example.

- (28) dva syn-y
two son-GEN.PL
'two sons'

In (28) the noun *sons* in the genitive case follows the numeral *dva* 'two' producing a sentence grammatical in accordance with Ukrainian grammar. However, there are some exceptions when it comes to the noun 'year' *hod*. The paradigm of this noun with numerals in Russian and Ukrainian is shown in the Table 4 below.

Table 4: Paradigm of the noun *year* in Russian and Ukrainian

Russian	Ukrainian
1 god.NOM.SG	1 rik.NOM.SG
2–4 goda.GEN.SG	2–4 roky.NOM.PL
5–20 let.GEN.PL	5–20 rokiv.GEN.PL

In Russian, the noun *year* in the nominative case combines either with the numeral 1 or with any compound numeral whose last part is 1, such as 21, 31. The numerals ending in 2, 3, or 4 require a noun in the genitive case. The numerals from 5 to 20 require the suppletive genitive plural form *let*. In Ukrainian, the noun *rik* in the nominative case combines either with the numeral 1 or a compound numeral ending in 1. The numerals ending in 2 – 4 require a noun in the Nominative plural. The numerals from 5 to 20 require genitive plural form. Surzhyk speakers use the noun ‘*hod*’ with numerals 5 to 20. Consider the following examples:

(29) tryt’siat’ p’iat’ **hod**
 thirty five year.NOM.SG
 ‘thirty five years’

(30) sem’siat’ **hod**
 seventy year.NOM.SG
 ‘seventy years’

Patterning of numerals with nouns in examples (29)-(30) does not follow either Russian or Ukrainian. For example, in (29) the DP ‘thirty five years’ should be combined with the noun *rokiv* if following the Ukrainian grammar or *let* if following the Russian. The almost identical form of Russian and Ukrainian numerals makes their analysis more difficult, as in (30) where the numerals could be interpreted as coming from either Russian or Ukrainian, although the phonological closeness with Ukrainian and usage of Ukrainian numerals throughout the data indicates that this NumP does come from Ukrainian.

Sociolinguistic characteristics of the speakers and the DPs containing numerals discussed above indicate that Surzhyk can hardly be a case of composite codeswitching. More data on various DPs with numerals are needed in order to analyze this phenomenon in greater depth. If there are more nouns that pattern similarly to the noun ‘year’, it will indicate that Surzhyk is a mixed language with the late system morphemes from Russian supplying agreement markers in one constituent; namely, the DP containing a numeral. On the other hand, the idiosyncrasy of the DPs shown above may simply be due to forms of the nouns ‘year’ and the way they decline. More research is needed to confirm or disprove these conjectures.

Lack of evidence of active Russian-Ukrainian bilingualism as well as inability of Surzhyk speakers to switch codes suggests that Surzhyk is sociolinguistically closer to a mixed language than to codeswitching. Recall the discussion of Auer’s continuum (1999) in which

is stigmatized.

References

- Auer, Peter. 1999. From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *The International Journal of Bilingualism* 3(4): 309-332.
- Bilaniuk, Laada. 2004. A typology of Surzhyk: Mixed Ukrainian-Russian language. *International Journal of Bilingualism* 8(4): 409-425.
- Bilaniuk, Laada. 2005. *Contested tongues: Language politics and cultural correction in Ukraine*. Ithaca: Cornell University Press.
- Bilaniuk, Laada & Svitlana Melnyk. 2008. A tense and shifting balance: Bilingualism and education in Ukraine. *International Journal of Bilingual Education and Bilingualism* 11(3-4): 340-372.
- Bilous, Tonia. 2005. Ukrainian IPA. <http://www.vesna.org.ua/txt/mov.html> (7 October, 2009).
- Flier, Michael. 1998. Surzhyk: The rules of engagement. *Harvard Ukrainian Studies* 22: 113-136.
- Flier, Michael. 2008. Surzhyk or Surzhyks? In Gerd Hentschel & Siarhiej Zaprudski (eds.), *Belarusian trasjanka and ukrainian surzhyk: Structural and social aspects of their description and categorization*, 39-56. Oldenburg: BIS-Verlag der Carl von Ossietzky Universitaet Oldenburg.
- Jones, Daniel and Dennis Ward. 1969. *The phonetics of Russian*. London: Cambridge University Press
- MacSwan, Jeff. 1999. *A minimalist approach to intrasentential code switching*. New York: Garland Publishing, Inc.
- Masenko, Larysa. 2008. Surzhyk: Istoriiia formuvannia, suchasnyĭ stan, perspektyvy funkcionuvannia. In Gerd Hentschel & Siarhiej Zaprudski (eds.), *Belarusian trasjanka and ukrainian surzhyk: Structural and social aspects of their description and categorization*, 39-56. Oldenburg: BIS-Verlag der Carl von Ossietzky Universitaet Oldenburg.
- Myers-Scotton, Carol. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. New York: Oxford University Press.
- Myers-Scotton, Carol. 2003. What lies beneath: Split (mixed languages) as contact phenomena. In Yaron Matras & Peter Bakker (eds.), *The mixed language debate*, 73-106. Berlin: Mouton de Gruyter.
- Myers-Scotton, Carol. 2006. *Multiple voices: An introduction to bilingualism*. Malden: Blackwell Publishing.
- Ohienko, Ivan. 2001. Istoriiia ukrainskoï literaturnoï movy. <http://www.litopys.org.ua/ohukr/ohu.htm> (22 September, 2007.)
- Ozerova, Nina (ed.). 2003. *Sopostavitelnaia grammatika russkogo i ukrainskogo iazykov*. Kyiv: Naukova Dumka.

Language Contact: Morphosyntactic Analysis of Surzhyk Spoken in Central Ukraine

- Ponomariv, Oleksandr (ed.). 2001. *Suchasna ukraïnska mova: Pidruchnyk dlia studentiv humanitarnykh spetsialnostei vyshchykh zakladiv osvity*, 2nd edn. Kyiv: Lybid.
- Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish y termino en Espanol: toward a typology of codeswitching. *Linguistics* 18: 581-618.
- Pugh, Stefan. 2007. *A new historical grammar of the East Slavic languages*. Muenchen: Lincom Europa.
- Pugh, Stefan & Ian Press. 1999. *Ukrainian: A comprehensive grammar*. London: Routledge.
- Serbenska, Oleksandra. 1994. *Antysurzhyk*. Lviv: Svit.
- Shevelov, George. 1977. On the chronology of *h* and the new *g* in Ukrainian. *Harvard Ukrainian Studies* 1: 137-152.
- Shevelov, IUri. 1998. *Ukraïns'ka mova v pershii polovyni dvadtsiatoho stolittia, 1900-1941: Stan i status*. Chernivtsi: Ruta.
- Shevelov, IUri. 2002. *Istorychna fonolohiia ukraïnskoï movy*. Kharkiv: Akta.
- Stavytska, Lesia & Volodymyr Trub. 2007. Surzhyk: Mif, mova, komunikatsia. In Lesia Stavytska (ed.), *Ukrainsko-rosiis'ka dvomovnist'*, 31-121 Kyiv: Pulsary.
- Thomason, Sarah & Terrence Kauffman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Tyshchenko, Kostiantyn. 2000. *Metateoriia movoznavstva*. Kyiv: Osnovy.
- Ushakov, Dmitrii (ed.). 1940. *Tolkovyï slovar' russkogo iazyka*. Moscow: Sovetskaia entsiklopediia.
- Vakhtin, Nikolai, Oksana Zhironkina, Yekaterina Romanova, & Irina Liskovets. 2003. Opre-delenie Surzhyka v sotsiolingvisticheskom aspekte. <http://www.eu.spb.ru/ethno/projects/project3/ukraine/007/009.htm> (15 November, 2007.)
- Zilynskyi, Ivan. 1979. *A phonetic description of the Ukrainian language*. Cambridge: Harvard Ukrainian Research Institute.

University of Minnesota – Twin Cities
Department of Linguistics
214 Nolte Center
315 Pillsbury Drive SE
Minneapolis, MN 55455

kent0082@umn.edu