



Universiteit
Leiden
The Netherlands

**CLASSIFYING UKRAINE'S NON-STANDARD LANGUAGE VARIETY:
Surzhyk as a Developing Mixed Language**
Colver, Julia

Citation

Colver, J. (2023). *CLASSIFYING UKRAINE'S NON-STANDARD LANGUAGE VARIETY::
Surzhyk as a Developing Mixed Language*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master thesis in the
Leiden University Student Repository](#)

Downloaded from: <https://hdl.handle.net/1887/3570968>

Note: To cite this publication please use the final published version (if applicable).

CLASSIFYING UKRAINE'S NON-STANDARD LANGUAGE VARIETY:

Surzhyk as a Developing Mixed Language

Julia Colver

MA Thesis Russian and Eurasian Studies, Faculty of Humanities, Leiden

University

Dr. Egbert Fortuin

7/7/2022

20305 Words

Preface	5
Current Events	5
Relevance	5
Transliteration	6
Methodology	6
Case Study Analysis	6
Overview	7
Topic	7
Pidgin Language Definitions	7
Creole Language Definitions	8
Mixed Language Definitions	8
Surzhyk Definitions	9
Research Question	10
Introduction	10
Historical Context	10
Contact Language Theories and Criteria	10
Preliminary Conclusions	12
Historical Background	13
Russification Policies	13
Promotion of Ukrainian	13
Effect on Surzhyk	14
Literature Review	14
Introduction	14
Common Features of Surzhyk	15
Regions of Surzhyk Speech	17
Past Research of Surzhyk	19
Stigma of Surzhyk	20
Definitions of Surzhyk	21
Surzhyk as Code-Switching	25
Surzhyk as a Pidgin Language	26
Surzhyk as a Mixed Language Variety	28
Research Gap	29
Conclusion	29
Can Surzhyk Even Be Considered a Contact Language?	30

Del Gaudio’s Theory of Surzhyk as a Sociolect	30
Surzhyk as a Contact Language	31
Pidgin Development	31
General Theories of Pidgin Development	31
Applying Theories of Pidgin Development to Surzhyk	34
Implications	35
Surzhyk and the Creole Continuum	35
The Concept of a Creole Continuum	35
Applying the Creole Continuum to Surzhyk	36
Implications	37
Surzhyk’s Lexifier Language	37
The Concept of a Lexifier Language	37
Applying The Concept of a Lexifier Language to Surzhyk	38
Implications	39
Mixed Language Development	39
General Theories of Mixed Language Development	39
Applying Theories of Mixed Language Development to Surzhyk	40
Implications	42
Complexity of Surzhyk	42
Contact Language Complexity	42
Applying Theories of Complexity to Surzhyk	43
Implications	45
Stability and Uniformity of Surzhyk	46
Contact Language Stability	46
Applying Theories of Stability to Surzhyk	48
Implications	51
Mixed Language Categories	52
Intertwined and Converted Mixed Languages	52
Surzhyk as an Intertwined Classical Mixed Language	54
Implications	55
Symbiotic Mixed Languages	55
Plain and Symbiotic Mixed Languages	55
Surzhyk as a Symbiotic Mixed Language	56
Figure 1: Self-Reported Native Language in Ukraine	57
Implications	58

Surzhyk as an Identity Marker	59
Contact Languages as Identity Markers	59
Surzhyk as a Form of In-Group Identity	59
Implications	61
Surzhyk as a Native Language	62
Contact Languages and Nativity	62
Can Surzhyk Be a Native Language?	63
Implications	64
Surzhyk as a Codified Linguistic System	65
Mixed Language Theory and ‘Fused Lects’	65
Surzhyk as a ‘Fused Lect’	66
Implications	66
Conclusion	67
Classification of Surzhyk	67
Future Research	69
Bibliography	71

Preface

Current Events

It is important to acknowledge the importance of current political events and their implications on this research. Recent events in the region have greatly changed the scope of this research project, meaning that the methodology and source material of this thesis had to be adapted in response to these limitations. Due to Russia's invasion of Ukraine, it was not possible to conduct field research on Surzhyk in Ukraine, as interviews and surveys could not be conducted due to both practical and moral limitations. As a result, in this thesis the focus is on theoretical research, using secondary literature on contact languages as well as secondary literature discussing Surzhyk in Ukraine.

Relevance

Additionally, the recent Russian invasion of Ukraine means that the research on this topic is increasingly relevant in the context of the Ukrainian linguistic landscape. As a result of Russia's invasion of Ukraine, the use of Surzhyk has become increasingly relevant in discussions about language identity in Ukraine. Ukrainian identity is increasingly being shaped by the war with Russia, and this may also shape attitudes towards Surzhyk as an expression of 'anti-Ukrainian' identity. The Russian invasion has greatly impacted language attitudes in Ukraine, particularly when it comes to the use of standard Russian or other varieties of non-standard Ukrainian. Because of this, there may be greater stigmatization of Surzhyk as a 'Russified' language variety, and Surzhyk may be increasingly perceived as something Russian and therefore foreign. More research is needed to determine the sociolinguistic effect of the conflict with Russia on Surzhyk, but it is clear that the research of Surzhyk is increasingly relevant as a result of current events in the region.

Transliteration

In the literature on this topic, there exists a great variety in transliteration styles. The word *суржик* itself has been transliterated differently by various authors who are writing on the subject, and there appears to be no universal standard that is used by all sources. Some authors choose to use the International Phonetic Alphabet, which would render *суржик* as 'Surʒɪk, while others choose to transliterate the word using the International Scholarly System, transliterating *суржик* as *Suržyk* or *Suržik*. Still others do not appear to follow a particular transliteration style and use the spelling *Surzhyk*. While all of these transliteration styles are found in the various sources discussing *суржик*, the spelling *Surzhyk* appears to be the one used most often. Therefore, for consistency's sake, the spelling *Surzhyk* will be used throughout this thesis.

Methodology

Case Study Analysis

The methodology used in this thesis is a case study analysis. Because of the limitations on conducting practical research, in this thesis theoretical research of secondary literature will be conducted in order to analyze Surzhyk as a language variety. Theories of contact language development and classification criteria to the case study of Surzhyk will be applied, using the literature on contact languages, in particular literature on pidgins, creoles, and mixed languages. Through this methodology, a classification of Surzhyk as a fused lect or developing mixed language will be suggested.

Overview

Topic

Суржик, or Surzhyk, is a linguistic ‘mix’ of Ukrainian and Russian that is spoken by about ten percent of the Ukrainian population. The term *Surzhyk* originally referred to a type of mix of wheat and rye, and gradually came to denote a mix of the Ukrainian and Russian languages which was spoken mainly in eastern and central Ukraine. This term had a negative connotation as representing an ‘impure’ and undesirable form of Russian or Ukrainian, and the stigma around the use of this variety still exists today. Generally, the main linguistic characteristic of Surzhyk is that it breaks linguistic norms, as it does not adhere to the rules of either the Ukrainian or the Russian standard language, while the main social characteristic of this language variety is that it has a low status (Bernsand 2001, 40). There are many varying definitions of Surzhyk, and there is little agreement among researchers as to how this language variety should be classified, whether as a form of code-switching, a pidgin, a creole, a mixed language, or as something else entirely.

Pidgin Language Definitions

For the purposes of this research, it is necessary to define several of the linguistic terms that will be used throughout this thesis. A pidgin is usually defined as a contact language that arises as a result of trade or other economic exchange between speakers of different languages. Bakker and Matras define a pidgin language as a language which functions as a lingua franca and which is lexically and structurally limited in its scope of communication (Bakker and Matras 2013, 25). According to this definition, pidgins are norm-based communication systems that have developed in situations where several groups created communication systems that are not fully expressive languages (Bakker and Matras 2013, 53). Kaye and Tosco additionally define

pidgins as a permanent social solution to a language barrier that involves the creation of a new language (Kaye and Tosco 2001, 28). As a result, a pidgin language generally develops in a multilingual setting where at least three languages are spoken, of which none are used as a common means of communication (Ibid). Therefore, pidgins are often viewed as simplified communication systems that combine elements from several different and unrelated languages, as they are generally considered to have developed as a lingua franca in trade situations where participants do not share a common language.

Creole Language Definitions

Following from this definition of pidgin languages, creoles are generally considered to be contact languages that develop from pidgin languages. The literature on contact languages defines a creole as a pidgin language which has evolved to contain native speakers (Kaye and Tosco 2001, 34). Pidgins and creoles are described as having one lexifier language, meaning that they derive the majority of their lexicon from that language (Bakker and Matras 2013, 65). Creoles are generally viewed as the next step in contact language development following the creation of a pidgin, and as such are seen as more complex languages that have moved beyond the context of economic exchange to be the primary or native language of a particular community, even if they develop and change more rapidly than ‘normal’ languages.

Mixed Language Definitions

The third type of contact language discussed in this thesis is a mixed language. A mixed language is usually defined as a language variety that arises in bilingual settings and combines elements from different language families or branches (Mazzoli and Sippola 2021, 1) and whose grammar and lexicon cannot be attributed to a single source language (Bakker and Matras 2003,

21). Mixed languages can resemble code-switching or sociolects of their source languages, (Bakker and Matras 2013, 200-201), and are sometimes argued to have developed as a result of code-switching (McConvell 2008, 187). However, they are generally considered to be autonomous linguistic systems in their own right (Viveka 2015, 70). Mixed languages are often described as having two clear parent languages, as they combine elements of two different languages (Bakker and Matras 2003, 180). As well as linguistic theories about pidgin and creole languages, theories about mixed languages offer useful insights when analyzing Surzhyk as a language variety.

Surzhyk Definitions

In addition to defining these three contact languages, it is necessary to explain the definition of Surzhyk as it will be referred to in this thesis. As will be discussed in the literature review, there are many varying definitions of Surzhyk in the existing literature. Surzhyk has been defined by different authors as a substandard language variety, a form of code-switching, a mixed language, a pidgin, or a nonstandard dialect of Ukrainian or Russian. The definition that will be referred to in this thesis is the broad definition of Surzhyk as a non-standard language variety that has Ukrainian as its primary linguistic base. Although this definition is by far not a comprehensive definition of all the variety that is present in different regional ‘Surzhyks’, for the purposes of this thesis this definition will be used in order to analyze and classify Surzhyk as a linguistic variety. In this thesis, these many definitions will be discussed and Surzhyk in the context of theories of contact language development will be analyzed. In doing so, classification of Surzhyk as a type of mixed language will be suggested.

Research Question

Therefore, the research question discussed in this thesis is the following: to what extent can Surzhyk be considered a pidgin language? How should Surzhyk best be classified?

Introduction

Historical Context

First, the historical background of linguistic policy in Soviet and modern-day Ukraine and how this has affected the stigma around the use of Surzhyk will be discussed. As a result of Russification policies, linguistic policy in present-day Ukraine has emphasized the use of the Ukrainian language, leading to a stigmatized view of non-standard linguistic varieties, such as Surzhyk. Whether Surzhyk can be seen as a contact language will be discussed, using theories and classification criteria found in the literature on contact languages. A tentative classification of Surzhyk as a fused lect or developing mixed language will be offered.

Contact Language Theories and Criteria

Firstly, theories of contact language development are useful to apply to the analysis of Surzhyk. Monogenetic theories of pidgin and creole development stress a common ancestor among different pidgins and ignore the role of the lexifier language, while Polygenetic theories emphasize the continuity between the target language and the pidgin or creole. Using these theories, one can apply Monogenetic approaches to Surzhyk and view the research on Surzhyk as emphasizing what different regional varieties of Surzhyk have in common. In addition, using the concepts of the creole continuum and lexifier languages, one can conclude that like fused lects, Surzhyk is perhaps on a continuum between Russian and Ukrainian. While Russian could be considered Surzhyk's lexifier language from which it acquires the majority of its lexicon, this

definition does not fully explain the structure of Surzhyk, as Surzhyk's vocabulary borrows from both its source languages to a greater extent than creoles do. Unidirectional and fusional theories of mixed language development are also useful in the analysis of Surzhyk. Unidirectional theories propose a one-way shift from the source language to the mixed language, while fusional theories propose that two source languages combine to form a third new language. Using these theories, one can compare Surzhyk to a fusional mixed language, demonstrating that it has more in common with mixed languages than other contact languages. In addition, two criteria of contact languages that are useful to this analysis are language complexity and language stability. These criteria are often used in the general literature on contact languages, and as such are useful in the classification of Surzhyk. While pidgins are considered to lack certain complexity and contain more linguistic variation, creoles and mixed languages are generally viewed as more complex and stable linguistic varieties that have greater speaker norms and linguistic conventions. Like mixed languages, Surzhyk appears to be a relatively stable and complex language variety. Therefore, it is useful to discuss some important criteria of mixed languages and apply them to Surzhyk in order to discuss if Surzhyk can be defined, albeit tentatively, as a mixed language. Intertwined mixed languages are composed of two mutually dependent components that form a unique whole, while converted mixed languages are languages that have adopted the grammar of another language without changing the lexicon. Using these criteria, one could view Surzhyk as a type of intertwined mixed language. Similarly, Surzhyk could be viewed as a symbiotic mixed language, meaning that it is a contact language that is commonly spoken alongside the source languages in the Surzhyk-speaking community. Mixed languages are generally considered to be identity markers, and the same could be said for Surzhyk, as it is a linguistic variety that is not spoken in response to a communication gap but rather for the

purposes of in-group identity. Language nativity is another important criteria that is discussed in the contact language literature. While pidgins are not considered to be native languages, creoles and mixed languages can be a speaker's native language. Applying this criteria to Surzhyk, we can conclude that like mixed languages, Surzhyk can be considered a speaker's native language in some contexts, although the degree to which one can be a 'native' speaker of Surzhyk varies considerably. Contact languages are also viewed as codified autonomous systems, an important distinction to apply to Surzhyk, as it is not a jargon but rather a fused lect with particular rules.

Preliminary Conclusions

Using these criteria discussed in the literature on contact languages, Surzhyk appears to share significant characteristics in common with both mixed languages and creoles. While the categories of pidgins, creoles, and mixed languages are merely overarching categories which are used to describe contact languages, and as such cannot perfectly describe or explain Surzhyk, analyzing their classification criteria in relation to Surzhyk nonetheless offers valuable insight to explain Surzhyk's unique linguistic features. There is some debate as to whether Surzhyk can be considered a true contact language, and some authors view it as more of a sociolect or dialect of Ukrainian. However, while contact language theory and classification criteria do not perfectly describe Surzhyk, these theories remain the most applicable to the analysis of Surzhyk conducted in this thesis, as Surzhyk shares significant similarities with contact languages such as creoles and mixed languages. Using the theories and classification criteria discussed in the general contact language literature, therefore, in this thesis a tentative classification of Surzhyk as a fused lect or developing mixed language is offered.

Historical Background

Russification Policies

Surzhyk is often viewed negatively by Ukrainian language purists as a form of broken Ukrainian. As a result, discussions of Surzhyk are often intertwined with discussions of linguistic ideology. It is thus useful to provide a brief historical background of linguistic policy in Ukraine and clarify how this pertains to the analysis of Surzhyk conducted in this thesis. The Russian and Ukrainian languages have long coexisted in Ukraine, and there is a relatively high degree of mutual bilingualism throughout the country (Tovares 2019, 463). Soviet policies of Russification have long favored the Russian language, granting Russian a privileged position until the independence of Ukraine (Taranenko 2007, 121).

Promotion of Ukrainian

Because of this earlier emphasis on Russian and Russification, Ukraine's language policies have increasingly emphasized the role of the Ukrainian language. This emphasis has only become stronger as a result of the annexation of Crimea in 2014 and increasing political tensions with Russia, meaning that there have been greater efforts by the government to promote the use of Ukrainian rather than Russian in most spheres of communication. There is still widespread bilingualism throughout the country, but since 2014 there has been greater emphasis on the use of the Ukrainian language in the government, education system, and daily life. One example of this is the *Закон України «Про забезпечення функціонування української мови як державної»*, or the Law of Ukraine 'To ensure the functioning of the Ukrainian language as a state language', which stipulates that the Ukrainian language must take priority in most spheres of public life, in particular in education, public administration, media, and advertising (Оснач 2019). Ukrainian language activists have emphasized the need for clearer boundaries between

the two languages (Bernsand 2001, 40), thereby further promoting the use of standard ‘official’ Ukrainian. The mainstream discourse in Ukraine often presents Ukrainian as an innocent or a martyr language, portraying it as the victim of a deliberate attack by the enemies of Ukraine (Sériot 49). This is particularly applicable to the current political context, as the war with Russia has caused many Ukrainians to view Russian as the language of the occupier (Solomon 2022).

Effect on Surzhyk

As a result of this promotion of Ukrainian and an increasing tendency to see Russian as the language of the occupier, there has arguably been greater stigmatization of Surzhyk, portraying it as a form of ‘impure’ Ukrainian. Surzhyk has increasingly become a symbol of ‘cultural lowness’ and representative of the damage done by Russification, meaning that nonstandard language forms were used in comedy or as part of the derogatory portrayal of less educated, provincial Ukrainians (Bilaniuk 2018, 296). According to this discourse, “the speaker of Surzhyk who has allowed his Ukrainian language to be corrupted by elements of a foreign language without mastering that language, represents a sort of cultural dead zone between the Ukrainian and Russian cultures and is by many nationalists not considered to be a full member of any of them” (Bernsand 2001, 43). This historical background is important to keep in mind, as Surzhyk is often viewed as an impure and corrupted variety of Ukrainian due to nationalist language ideologies and the increasing emphasis on speaking Ukrainian instead of Russian.

Literature Review

Introduction

Most authors have focused on theoretical lexical analysis, but there is a lack of research when it comes to analyzing the grammar or development of Surzhyk. Even the definition of the

term *Surzhyk* itself is contested, and many authors use this term to refer to different linguistic phenomena. Surzhyk has been defined in the existing literature as a broad overarching term for non-standard language use, as a language variety of Ukrainian that contains Russian elements, as a language variety of Russian that contains Ukrainian elements, and as a social dialect. In addition, there is little agreement on how Surzhyk should be classified. Some authors argue that Surzhyk should be classified as a form of code-switching, others view it as a pidgin or creole language, and others argue it should be classified as a supra-dialectal mixed language or as an entirely new linguistic phenomenon. There is much disagreement among different authors about the definition and classification of Surzhyk, and this lack of consensus in the existing literature further complicates the research of this language variety. Additionally, much of the existing research on the subject of Surzhyk is outdated, and negative views of Surzhyk still persist in the existing Russian-language literature. As a result, much of the existing literature is not relevant to the context of the contemporary linguistic landscape in Ukraine. This research paper will attempt to address this gap in the literature by analyzing the linguistic structure and the ongoing development of Surzhyk in the context of contact language development and criteria.

Common Features of Surzhyk

Most existing literature has focused on theoretical lexical analysis, but there is far less research on the grammar or sociolinguistic development of Surzhyk. However, there is some consensus among authors on the pronunciation and linguistic structure of Surzhyk. It is unclear how Surzhyk is pronounced, but most authors agree that the phonetics of Surzhyk are closer to the pronunciation of standard Ukrainian than standard Russian (Groen 2014, 203). Generally, Surzhyk speech appears to exhibit an adaptation of Russian lexemes to Ukrainian phonology (Del Gaudio 2010, 65). According to Del Gaudio, Surzhyk generally demonstrates six vocalic

phonemes, namely i, y, e, a, o, and u (Del Gaudio 2010, 66). Similar to the pronunciation of Ukrainian, in Surzhyk the Russian -g is pronounced as a laryngeal -h (Romanova, Zhironkina, and Vakhtin 2007, 332). Similarly, in Surzhyk hard consonants are generally used in Russian loan words, and consonants are generally soft before -e and -i, as in Ukrainian (Groen 2014, 205). In general the phenomenon known as *akan'e*, in which the unstressed vowel -o is pronounced as -a, is absent in Surzhyk. Surzhyk words, whether they originate from Russian or Ukrainian, do not reflect the Russian pronunciation of the unstressed vowel -o as -a (Курохтина 2009, 106). According to Del Gaudio, this unstressed -o does, however, emerge in “words of clear Russian origin, when Ukrainian dialectisms and archaisms happen to be morphologically similar to standard Russian or when the interlocutor speaks in this language and the speech fundamentally tends towards Russian” (Del Gaudio 2010, 66-67). This lack of *akan'e* in Surzhyk speech resembles the Ukrainian pronunciation of an unstressed -o, which is pronounced as a long -o rather than -a. In addition to the Ukrainian pronunciation of the unstressed -o, Surzhyk includes the use of the Ukrainian -i vowel (Nazarenko 2012, 245). Therefore, most authors agree that Surzhyk is generally pronounced according to Ukrainian phonetics, although there is a clearly observable Russian phonetic influence. For example, in Surzhyk, the Ukrainian -i is often replaced by the Russian -o, as in *ножки* versus *ніжки* and *боль* versus *біль* (Del Gaudio 2010, 68). However, there are also some features of pronunciation in surzyk that do not adhere to either Ukrainian or Russian pronunciation rules (Gasparov 2006, 123). According to Del Gaudio, some of Surzhyk's more idiosyncratic traits are linguistic features such as the reduced presence of the Ukrainian *є* sound at the beginning of a syllable or the substitution of the Ukrainian sounds *дж* and *джэ* with a pronunciation similar to *з* and *жэ* (Del Gaudio 2010, 69-72). As a result, according to Del Gaudio, a main distinguishing feature of Surzhyk, in addition to its lexicon and

phonetics, is its morphology, or the structure of its vocabulary (Del Gaudio 2010, 73). Similarly, although there is less research on the grammar of Surzhyk, most authors writing on the subject agree that it does not appear to deviate significantly from Ukrainian grammar. According to Курохтина, the grammatical forms of Surzhyk correspond to the norms of Ukrainian, and forms that correspond to the norms of the Russian language are an exception. (Курохтина 2009, 109). Authors discussing Surzhyk appear to agree that most of the grammar, syntactic and intonational constructions, as well as phonetics of Surzhyk come from Ukrainian, while much of the vocabulary is Russian in origin (Кузнецова, Савченко, Хмелевский 2018, 95). However, authors such as Del Gaudio argue that the “very basis of Surzhyk...is undoubtedly Ukrainian” (Del Gaudio 2010, 88). According to this view, Surzhyk is a primarily Ukrainian variety with a particular Russian influence, although it varies from region to region. There are some differences in regional use of Surzhyk, and some authors argue that several different idiolects of Surzhyk are used in different areas of Ukraine, although they argue that these varieties all generally follow particular linguistic rules (Stadnik-Holzer 2007, 367). For example, Gasparov argues that the eastern version of Surzhyk has more features in common with Russian than the version of Surzhyk that is spoken in central Ukraine (Gasparov 2006, 122). Most of the existing research has focused on lexical analysis, and although there is a lack of research on the grammar of Surzhyk, these authors generally agree that many linguistic features of Surzhyk are closest to Ukrainian while much of the vocabulary is Russian.

Regions of Surzhyk Speech

Surzhyk is spoken in several different areas of Ukraine, in both the east and the center of the country, and the existing literature generally agrees that it is spoken by different social groups. According to Romanova, Zhironkina, and Vakhtin, Surzhyk is spoken both by those who

live and work most of the time outside large cities with low social mobility and by those who live or work in a large city with higher social mobility (Romanova, Zhironkina, and Vakhtin 2007, 339). There is some debate among different authors on how Surzhyk originated. Del Gaudio argues that Surzhyk originated in factories and mines in Donbas, among poorly educated residents and military personnel (del Gaudio 2015, 223). Nazarenko, however, argues that Surzhyk appeared in the eighteenth century when Russian became the language of state administration in Ukraine. Despite this, she agrees with del Gaudio that the first speakers of this variety were likely members of the army stationed in Ukraine as well as local officials (Nazarenko 2012, 245-246). Another view is that of Кузнецова, Савченко, and Хмелевский, who argue that another type of Surzhyk originated in Kiev during the eighteenth and nineteenth centuries (Кузнецова, Савченко, and Хмелевский 2018, 99). Several authors argue that Surzhyk is the primary language for some Ukrainians as a form of daily communication within the family. Despite the low status of this variety, Surzhyk is often described in the literature as the language of everyday household and family communication (Nazarenko 2012, 249). Surzhyk is generally considered by different authors to be a primarily spoken language with no real written form. (Del Gaudio 2010, 73), in part because of its low status as a non-standard language variety. Although survey responses may reflect the view that the state language should be used in all spheres of life, Surzhyk is often used in everyday communication (Кузнецова, Савченко, and Хмелевский 2018, 94). However, because of the stigmatization of this language variety, when asked about their use of Surzhyk, many people may deny speaking this variety and claim that they speak Russian or Ukrainian (Fomina 2014, 5). In addition, Surzhyk is also used in the media to convey humor or satire. Surzhyk is often used in popular culture in order to accommodate “the linguistic tastes of the ‘enlightened’ audience who want to be amused without being disturbed or

offended” (Gasparov 2006, 119). Authors such as Gasparov and Nazarenko agree that Surzhyk has long been used in Ukrainian literature and theater as a means of creating comic effect (Nazarenko 2012, 249). By portraying Surzhyk as the language of an uneducated, rural population, Ukrainian popular culture has long used this language variety to achieve a humorous effect. Therefore, it appears that Surzhyk has been widely used since the eighteenth century in both daily life and popular culture, but it has been under-researched as a result of the stigma associated with speaking this variety.

Past Research of Surzhyk

During the Soviet period, Surzhyk was ignored by Soviet linguists and researchers. Researchers at the time ignored this nonstandard form of speech because its existence contradicted Soviet ideals, which stressed the development of national languages (Nazarenko 2012, 247). As a result, the linguistic study of Surzhyk as a linguistic phenomenon began only after the collapse of the USSR and the creation of an independent Ukraine (Nazarenko 2012, 247). Because of this ignorance by Soviet linguists, most existing research on Surzhyk has been conducted during the late 1990s and early 2000s. There is little existing research that was conducted after 2014, meaning that the literature does not take into account the changing linguistic situation in Ukraine since the annexation of Crimea and the conflict in Donbas and Luhansk. In addition, much of this existing research has focused on the linguistic structure of Surzhyk rather than its development in Ukrainian society. As a result, there is a lack of recent research on the development and use of this language variety in Ukraine today.

Stigma of Surzhyk

Surzhyk is often seen as a stigmatized language variety with little social prestige. According to Groen, Surzhyk and its speakers are associated with low status (Groen 2014, 192). Surzhyk, along with other nonstandard language use in Ukraine, lacks any official recognition or institutional power (Groen 2014, 192), and many researchers continue to describe it as a negative and undesirable language variety. In much of the Russian-language literature on Surzhyk, this variety is viewed as a ‘lowly phenomenon’ associated with low prestige and negative connotations, and as a distortion of both Ukrainian and Russian spoken by the poorly educated (Gasparov 2006, 117). A negative attitude towards Surzhyk and its speakers has resulted in a lack of recent and detailed research on this language variety (Groen 2014, 194). This negative view of Surzhyk is predominantly found in Russian-language works by both Russian and Ukrainian authors. Taranenko argues that elements of the Russian language have polluted the structure of the Ukrainian language, resulting in the use of Surzhyk (Taranenko 2007, 124-125). This perspective of Surzhyk as a lesser variety of Ukrainian is shared by Shevchuk, who argues that it is a substandard variety of Ukrainian. According to this perspective, Surzhyk is a kind of Ukrainian vernacular that was formed as a result of the undesirable merging of Russian and Ukrainian (Shevchuk 2021, 359). This negative view of Surzhyk argues that this language variety was created without any norms and, and it stipulates that it poses a threat to Ukrainian as the national language (Shevchuk 2021, 359). Authors such as Shevchuk and Taranenko view Surzhyk as a lesser variety that is damaging to the development of standard Russian and Ukrainian. Shevchuk, for example, argues that Surzhyk blurs the stylistic boundaries of journalistic language, threatening standard speech as Russian and Ukrainian speakers begin to use ‘street language’ (Shevchuk 2021, 359). Similarly, Nazarenko argues that Surzhyk is a distorted form of the Ukrainian language, which includes borrowed elements of the Russian

language (Nazarenko 2012, 244-245). Like other Russian and Ukrainian authors, she views Surzhyk as a negative variety that threatens the purity of both standard Ukrainian and standard Russian. Nazarenko argues that in order to combat this threat, it is necessary to increase the public prestige of the standard Ukrainian language and promote the use of standard Ukrainian in informal spheres of communication (Nazarenko 2012, 250). In her view, this will lead to Ukrainians choosing Ukrainian as their main means of everyday communication, thereby eliminating the use of Surzhyk (Nazarenko 2012, 250). As a result of the negative view of Surzhyk in many Russian-language publications, much of the existing research on Surzhyk has not incorporated practical research on the development of this language variety in contemporary Ukrainian society, instead focusing on the need to promote the use of standard Russian.

Definitions of Surzhyk

The definition of the term *Surzhyk* itself is contested. There is no single overarching definition of this linguistic expression, and different authors may use the term *Surzhyk* to refer to different language varieties. Unclear and varying definitions further complicate conducting research on this topic, as there is little agreement on what language use actually constitutes Surzhyk. The term Surzhyk may refer to a wide variety of linguistic forms, and there is little agreement on when non-standard language use becomes Surzhyk (Groen 2014, 192). Many of the definitions that are given by different authors lack specificity, and many definitions attach a certain stigma to this language variety (Romanova, Zhironkina, and Vakhtin 2007, 327). This makes it difficult to determine what Surzhyk is, and it further complicates researching the development of this language variety. Different authors use the term Surzhyk to refer to different speech varieties that appeared as a result of language contact (Romanova, Zhironkina, and Vakhtin 2007, 329). One definition of Surzhyk that is used in the existing literature defines it as a

broad term for anything that is a non-standard form of Ukrainian or Russian. According to Groen, Surzhyk is often used as a “hyperonymic for everything that is not ‘pure’ Russian or Ukrainian”, meaning that dialects spoken in Ukraine may also potentially be labeled as Surzhyk (Groen 2014, 196). Groen further differentiates between three different categories of speakers and corresponding varieties of Surzhyk (Groen 2014, 195). He differentiates between the spontaneous use of Surzhyk, the conscious use of Surzhyk, and spoken Russian or Ukrainian that has elements of the other language because the speaker only speaks one of these languages well (Groen 2014, 195). He describes the spontaneous use of Surzhyk as a local Ukrainian dialect with frequent Russianisms that is spoken by people who do not speak either Ukrainian nor Russian sufficiently. In contrast, he explains the conscious use of Surzhyk as a variety spoken by those who speak both languages but not fluently. Finally, he argues that a third category of Surzhyk is the use of elements of one language in the other by those who only speak either Ukrainian or Russian fluently (Groen 2014, 195). Similarly, Bernsand argues that implicit discourses in Ukraine generally construct Surzhyk as a “monolingual code of speakers not commanding standard Ukrainian or Russian”, with extra-linguistic values and practices different from those of Ukrainian and Russian-speakers (Bernsand 2006, 111). Other authors may use the term Surzhyk to refer to more traditional code-switching, and this is further complicated by the fact that some people may report using Surzhyk when their language use can better be described as intersentential or intrasentential code-switching, namely code-switching between sentences or code-switching within sentences (Lakhtikova 2017, 158). A second definition that can be found in the literature discussing Surzhyk defines it as a language variety of Ukrainian that contains Russian elements. Kent argues that Surzhyk is often used to indicate a Ukrainian language variety in which the grammar is influenced by Russian elements which are not present in

standard Ukrainian (Kent 2010, 33). Del Gaudio similarly argues that Surzhyk refers to anomalies of Ukrainian speech that are influenced by Russian grammar or pronunciation (del Gaudio 2015, 237). This definition of Surzhyk classifies it as a variety of Ukrainian that contains Russian lexical items (Kent 2010, 39), and views it as a primarily Ukrainian language variety rather than a true mix of the Russian and Ukrainian languages. A third definition found in the literature, however, refutes this and classifies Surzhyk as not only a variety of Ukrainian but also as a language variety of Russian that contains Ukrainian elements. Some authors argue that both types of Surzhyk exist, namely Ukrainian-Russian Surzhyk and Russian-Ukrainian Surzhyk and distinguish these two as two different language varieties (Groen 2014, 196). Долешаль, Дубичниский, and Ройтер argue that there are two fundamentally different varieties of Surzhyk, namely a Ukrainian language variety with Russian elements and a Surzhyk based on the Russian language with Ukrainian elements (Долешаль, Дубичниский, and Ройтер 2011, 249-250). Although del Gaudio refutes this idea of Surzhyk as a variety of Russian, he too acknowledges that some nonstandard Russian language use can potentially be included in his definition of Surzhyk (del Gaudio 2015, 237). In addition to these definitions, a fourth definition classifies Surzhyk as an intermediate language that emerged during a process of linguistic assimilation. Кузнецова, Савченко, and Хмелевский argue that Surzhyk should be viewed as an intermediate language that will be used until Russian displaces Ukrainian. They argue that Surzhyk is a hybrid of both Russian and Ukrainian that is used during a transition to using the Russian language while speakers are not actively fluent in either Russian or Ukrainian (Кузнецова, Савченко, and Хмелевский 2018, 94). Bilaniuk further defines five sub-categories under the term *Surzhyk*, namely urbanized peasant Surzhyk, village dialect-Surzhyk, Sovietized-Ukrainian Surzhyk, urban bilinguals' Surzhyk, and post-independence Surzhyk (Bilaniuk 2004, 409). According to

this categorization, urbanized peasant Surzhyk is a variety of Russian spoken by Ukrainians who lacked considerable proficiency in the language. She describes this type of Surzhyk as the ‘archetypal Surzhyk’ which resulted from industrialization and urbanization, as rural Ukrainian speakers came into contact with Russian-speaking administrators and moved to larger cities where Russian was the more prestigious language (Bilaniuk 2004, 416). Similar to urbanized peasant Surzhyk, Bilaniuk describes village-dialect Surzhyk as the Russification of Ukrainian (Bilaniuk 2004, 417). According to Bilaniuk, sovietized-Ukrainian Surzhyk is a language variety that resulted from Soviet Russification policies, making it an “institutionally created Fused Lect, a result of the Stalinist and later Soviet projects of eradicating national differences” (Bilaniuk 2004, 418). Urban bilinguals’ Surzhyk is described as a form of habitual language mixing by bilinguals spoken by urban Ukrainians in both the Soviet and post-Soviet period. Bilaniuk describes this form of Surzhyk as an irregular variety that lacks a common grammar, and that can take the form of code-switching or language mixing (Bilaniuk 2004, 419). The final category, post-independence Surzhyk, is described as a form of language mixing that resulted from the new state status of the Ukrainian language, as adults who were not accustomed to speaking Ukrainian tried to do so in the post-Soviet period (Bilaniuk 2004, 421). Similar to Кузнецова, Савченко, and Хмелевский, Bilaniuk’s outlines several different types of Surzhyk that were spoken as a result of a language transition to either Russian or Ukrainian. These categories by Bilaniuk define Surzhyk as an intermediate linguistic stage of sorts that is used in this phase of transition, or as an intermediate form of speech used by bilinguals. Finally, a fifth definition defines Surzhyk as a social dialect, arguing that it is closer to a linguistic jargon than a true mix of Ukrainian and Russian (Romanova, Zhironkina, and Vakhtin 2007, 328). This definition views Surzhyk as a particular social dialect of Ukrainian that is spoken as a result of frequent

contact with the Russian language. Because of these varying definitions of Surzhyk, researchers disagree on whether there is only one type of Surzhyk or whether different language varieties can all be considered Surzhyk. Some authors agree that there is only one linguistic variety called Surzhyk, but they disagree on the definition of this variety, and others argue that different forms of language use can all be considered Surzhyk (Kent 2010, 38). As a result of these different and contested definitions of Surzhyk, it is unclear what constitutes Surzhyk, making it unclear what different authors refer to when they discuss this language variety.

Surzhyk as Code-Switching

As well as disagreeing on the definition of Surzhyk, many authors disagree on how to classify this variety. Some authors argue that Surzhyk is simply another form of code-switching, while others reject this argument and claim that Surzhyk cannot be code-switching because it is a completely different linguistic phenomenon. Code-switching is an important phenomenon in the contemporary linguistic landscape in Ukraine. According to Groen, code-switching occurs in many different aspects of daily life in Ukraine, as people switch between Ukrainian and Russian frequently within a single conversation (Groen 2014, 190). As a result, Surzhyk is often perceived as a phenomenon that is close to or a form of code-switching. Kent argues that Surzhyk constitutes a form of code-switching, namely composite code-switching. This perspective argues that Surzhyk is a case of composite code-switching because “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” (Kent 2010, 45). Kent argues that because the motion verbs in Surzhyk do not have selected Russian prepositions, it is not a mixed language but rather a case of composite code-switching (Kent 2010, 45). She acknowledges that the existence of linguistic forms distinct from both Russian and Ukrainian might support the

classification of Surzhyk as a mixed language (Kent 2010, 49), but she argues that it is best described as a form of composite code-switching. She views it as composite rather than ‘traditional’ code-switching because the use of prepositions with verbs of motion in Surzhyk is due to structural changes in motion verbs influenced by Russian grammar rather than inter-sentential or intra-sentential code-switching between Russian and Ukrainian (Kent 2010, 47-48). Because Surzhyk speakers do not always display a bilingual command of both Ukrainian and Russian, Kent argues that it is not truly traditional code-switching but a composite variety in which speakers of Surzhyk combine grammar structures from both languages (Kent 2010, 51). This perspective views Surzhyk as a form of code-switching instead of a mixed language. Authors such as Kent acknowledge the unique structure of Surzhyk that makes it different from inter-sentential or intra-sentential code-switching, but argue that Surzhyk should be considered a form of composite code-switching rather than classifying it as a mixed language.

Surzhyk as a Pidgin Language

Authors discussing Surzhyk also disagree on whether it can be considered a pidgin or creole language. Barkovich is one of a group of authors who argues that Surzhyk could potentially be considered a pidgin language (Barkovich 2019, 274). According to this perspective, Surzhyk has no linguistic features caused by ‘related cross-lingual interference’ and should be classified as a pidgin or a creole language because it is a unique language variety but not a completely new or mixed language (Barkovich 2019, 282). Other authors, however, argue that Surzhyk cannot be considered a pidgin language because the origin of pidgin languages is different from that of Surzhyk. Groen argues that Surzhyk cannot be called a pidgin because a pidgin is the result of trade relations between two groups who speak completely different languages (Groen 2014, 192). He argues that a pidgin could not have resulted from language

contact between two strongly related languages, and argues that pidgins generally remain in an elementary stage of development that prevents full communication (Groen 2014, 192).

Additionally, Groen argues that Surzhyk also cannot be classed as a creole language because a creole is a language that results from a more developed pidgin (Groen 2014, 193). This perspective argues that Surzhyk allows for a far larger range of communication than is usually observed in pidgin languages, and that it therefore should not be considered a pidgin. On the other hand, Romanova, Zhironkina, and Vakhtin acknowledge that Surzhyk can potentially be defined as a pidgin language because like other pidgins, Surzhyk has based its lexicon on the language of the colonizer while keeping native pronunciation and simplifying morphology and syntax (Romanova, Zhironkina, and Vakhtin 2007, 328). However, like Groen, they argue that Surzhyk has a much lower extent of simplification than is usually observed in pidgins.

According to this perspective, Surzhyk allows a high degree of variation and can be a native language for its speakers, making it much more complex than a pidgin language. In addition, Romanova, Zhironkina, and Vakhtin argue that Surzhyk has not developed as a result of impeded communication because Russian and Ukrainian are mutually understandable to some extent (Romanova, Zhironkina, and Vakhtin 2007, 328). Similarly, Nazarenko argues that pidgins were formed as a result of interethnic contacts between Europeans and colonized peoples (Nazarenko 2012, 248). Like Groen and Romanova, Zhironkina, and Vakhtin, she argues that pidgin languages are much more simplified than Surzhyk and have much smaller vocabularies, as they remain only a means of interethnic communication (Nazarenko 2012, 248). This perspective argues that Surzhyk was the result of contact between two closely related languages, making it a mixed sublanguage rather than a pidgin or creole (Nazarenko 2012, 248). Therefore, although authors such as Barkovich and Nazarenko argue that Surzhyk can potentially be considered a

pidgin language, the perspective of Groen and others is that Surzhyk cannot realistically be considered a pidgin because it is far more complex and developed differently than pidgin languages.

Surzhyk as a Mixed Language Variety

In addition to the perspectives that Surzhyk should be considered a form of code-switching or a pidgin language, there is a perspective that views it as an entirely different linguistic phenomenon. Some authors who have this perspective view Surzhyk as a mixed language variety. Зинкевич views it as a supra-dialectal mixed language variety (Зинкевич 2013, 130), arguing that this definition better classifies it as a unique language variety different from the phenomena of code-switching and pidgin languages. Similarly, Курохтина views Surzhyk as a mixed language variety because its speakers are ‘monolingual’ and unable to switch to either Ukrainian or Russian (Курохтина 2009, 110-111). According to this perspective, insufficient knowledge of the norms of both Russian and Ukrainian prevents Surzhyk speakers from freely switching to either of these languages (Курохтина 2009, 110-111). As a result, Курохтина and others argue that because Surzhyk speakers are effectively monolingual, this language variety cannot be classified as either code-switching or as a pidgin language, but rather as a supra-dialectal mixed language variety. Yet other authors view Surzhyk as an entirely new linguistic phenomenon. Gasparov argues that although it originated from Ukrainian and Russian, Surzhyk is a “qualitatively new and distinct linguistic phenomenon” (Gasparov 2006, 126). This perspective views Surzhyk as a new linguistic phenomenon separate from code-switching and pidgin languages and argues that Surzhyk should be considered as something completely unrelated to these phenomena.

Research Gap

Most existing literature on Surzhyk has focused on theoretical lexical analysis, and there is far less research on the grammar or sociolinguistic development of Surzhyk. These authors disagree on the definition and classifications of Surzhyk, with some viewing it as a form of code-switching, others classifying it as a pidgin language, and still others arguing it is a supra-dialectal mixed language or a completely new and undefined linguistic phenomenon. In addition, many current studies of Surzhyk contain implicit negative biases towards this variety, meaning that there is a lack of practical research in this field. Much of the research on this topic is no longer fully relevant to the current linguistic situation in Ukraine, as much of the literature was published before 2014. The linguistic status of Ukrainian and Russian may have changed as a result of events such as the conflict in Donbas and the annexation of Crimea, meaning that more recent research on the development of Surzhyk in contemporary Ukraine is necessary. Because of negative views towards Surzhyk and because of the contested definition of this variety, many of the authors who have written on the subject have done little field work in Ukraine, meaning that in addition to being outdated, the existing literature is mostly theoretical work focused on the linguistic structure of Surzhyk. Therefore, this research will attempt to address the gap in the literature by analyzing the ongoing development of Surzhyk in the context of contemporary Ukrainian society.

Conclusion

Surzhyk is spoken in eastern and central Ukraine, and it is generally thought to have originated in the eighteenth century. Most existing literature has focused on theoretical lexical analysis rather than the grammar or the sociolinguistic development of Surzhyk, however, the existing literature does indicate that the phonology and grammar of Surzhyk appear closest to

standard Ukrainian while its lexicon appears to be closer to standard Russian. Surzhyk is a stigmatized language variety that is seen as a lesser form of Russian or Ukrainian, and this view is reflected in much of the existing Russian-language literature. In addition, the definition of the term *Surzhyk* itself is contested, as many authors use this term to refer to a broad range of linguistic phenomena. Surzhyk has been defined in the existing literature as a broad overarching term for non-standard language use, as a language variety of Ukrainian that contains Russian elements, as a language variety of Russian that contains Ukrainian elements, and as a social dialect. In addition, there is little agreement on how Surzhyk should be classified. Some authors argue that Surzhyk should be classified as a form of code-switching, others view it as a pidgin or creole language, and yet others argue it should be considered a supra-dialectal mixed language or even an entirely new linguistic phenomenon. There is a lack of consensus in the existing literature in defining and classifying Surzhyk, and this further complicates the research of this language variety. Much of the existing research is outdated and reflects a negative view of Surzhyk. Therefore, this research will attempt to address the gap in the literature by conducting more relevant research and by analyzing the ongoing linguistic development of Surzhyk.

Can Surzhyk Even Be Considered a Contact Language?

Del Gaudio's Theory of Surzhyk as a Sociolect

This thesis operates under the assumption that Surzhyk is a contact language and it applies general theories of contact language studies to this particular case study. While the general consensus among researchers is that Surzhyk should be considered a type of contact language, there are a few researchers who disagree, most notably Del Gaudio. Although this paper argues that Surzhyk can best be classified as a fused lect or developing mixed language,

the perspective of Del Gaudio is worth acknowledging. Del Gaudio posits that Surzhyk is not a result of language contact between Russian and Ukrainian but rather a ‘Russified’ Ukrainian dialect. According to this argument, there are many Ukrainian grammatical and lexical forms in Surzhyk that can be explained by a dialectological perspective of Surzhyk (Del Gaudio 2010, 139). Del Gaudio offers an alternative explanation for Surzhyk’s development that does not follow theories of contact language development, arguing that the formation of Surzhyk resulted from the combination of Ukrainian dialectal features (Del Gaudio 2010, 140). According to this view of Surzhyk, it is not a pure contact language but rather a language variety “based on dialectal interaction with the dominating languages” (Del Gaudio 2010, 230).

Surzhyk as a Contact Language

However, while Del Gaudio’s perspective offers interesting insights, following the analysis conducted in this thesis, Surzhyk seems to be best defined as a form of a contact language. Contact language categories do not fully explain Surzhyk, as they are merely generalized classifications, but Surzhyk can be observed to share significant similarities with contact language varieties such as mixed languages. Although every author seems to have a different definition of Surzhyk, meaning there is a great deal of disagreement on how to explain or characterize it, in the scope of this thesis contact language theories are more useful in analyzing Surzhyk than the dialectal theories used by Del Gaudio.

Pidgin Development

General Theories of Pidgin Development

Although Surzhyk is arguably not a pidgin, it is useful to include a description of theories about pidgin and creole development and apply them to this analysis of Surzhyk. Theories of

pidgin and creole development can offer some insight into the wider development and structure of Surzhyk, as there are some similarities between these theories and the development of Surzhyk. In particular, there are several theories that attempt to explain the development of pidgin and creole languages which are relevant for the analysis of Surzhyk conducted in this thesis. Generally, pidgins are seen as languages that result from two or more unrelated languages coming into contact, one of which is a superstrate, meaning that it has superior social status or prestige over the other languages (Kaye and Tosco 2001, 36). According to this widely-held view, the most important condition that must be met for the creation of a pidgin is that a common new language is needed (Kaye and Tosco 2001, 17). Most researchers of contact languages agree that a pidgin is an “interethnic contact language which is created when two or more ethnolinguistic groups lack any other common means of communication, but still have reasons to deal verbally with one another” (Bakker and Matras 2013, 20). Therefore, pidgins generally arise because these groups lack a common language with which to communicate. As a result, pidgins are often considered to be precursors to creole languages (Bakker and Matras 2013, 25), making them a more simplistic intermediate stage in the process of the creation of a creole language. According to this common argument, the development of pidgins as well as creoles arise as a result of interethnic trade exchanges. Bakker and Matras describe a ‘pidgin-creole life cycle’, in which there is first a type of linguistic jargon, then a pidgin, and then a creole (Bakker and Matras 2013, 25). This theory of a ‘pidgin-creole life cycle’ views pidgin development as a step in the broader development of creole languages. In connection to the common argument among researchers of contact languages that pidgins develop in a context where none of the various ethnolinguistic groups share a common language, there are two general types of theories that attempt to explain pidgin development, namely Monogenetic and Polygenetic approaches.

Monogenetic approaches emphasize what pidgins and creoles have in common and minimize the role of the lexifier, or target, language (Kaye and Tosco 2001, 55). Several different Monogenetic approaches to pidgin development are the Nautical-Jargon Theory, the Pidgin Portuguese Hypothesis (Kaye and Tosco 2001, 56-57), the foreigner talk theory (Kaye and Tosco 2001, 72-73), and the bioprogram hypothesis (ibid). The Nautical-Jargon Theory “ascribes to various forms of ‘nautical speech’ or ‘jargon’ ...the diffusion of certain common traits and lexemes across many different and widely-dispersed pidgins and creoles” (Kaye and Tosco 2001, 56). The Pidgin Portuguese Hypothesis argues that European-based pidgins and creoles are derived from a common grammar system which arose from a form of Pidgin Portuguese which developed in West Africa in the fifteenth century (Kaye and Tosco 2001, 57). The foreigner talk theory argues that pidgins develop as a result of the conscious simplification of one’s own language when used with foreigners who are assumed to lack a full understanding of the language (Kaye and Tosco 2001, 72), while the bioprogram hypothesis proposes that pidgins become creoles as a result of children acquiring a pidgin as their native language (Kaye and Tosco 2001, 73). All of these Monogenetic approaches have in common that they stress a common feature or ancestor among different pidgins, while downplaying the role of the lexifier language. In contrast to Monogenetic approaches, Polygenetic approaches emphasize the continuity between the target language and the creole, focusing less on what links different creoles (Kaye and Tosco 2001, 55). As a result, this approach minimizes continuity between pidgins and creoles (Ibid). This directly contradicts the ‘pidgin-creole life cycle’ theory, as it emphasizes the connection between the target language and the creole as being on a scale of proximity to the target or standard language while dismissing the development of creoles as resulting from pidgin development. Two main Polygenetic approaches to pidgin development are the view from below or the role of the

substratum theory and the view from above or the historical approach (Kaye and Tosco 2001, 59-64). The role of the substratum theory has two sides, namely the substratists and the universalists. The substratists attributed the developments of various pidgins and creoles to the influence of the substrate language, while the universalists explained this development by referring to ‘universal’ laws of linguistic evolution (Kaye and Tosco 2001, 59-60). In the view from above or the historical approach “any special link between pidgins and creoles is negated, and the indigenous contributions as well as the universal tendencies of creole development are radically downplayed” (Kaye and Tosco 2001, 64). These Polygenetic approaches both emphasize the connections between the target language and the pidgin or creole while ignoring similarities between different pidgins and creoles. Although it seems that Surzhyk cannot be classified as a pidgin, as will be discussed in later paragraphs, it remains useful to apply these theories of pidgin development to this analysis in order to come to preliminary conclusions about Surzhyk’s predominant features.

Applying Theories of Pidgin Development to Surzhyk

Using this literature on the development of pidgin languages, we can view the literature on Surzhyk as belonging to a Monogenetic approach rather than a Polygenetic approach. Authors discussing Surzhyk generally seek to emphasize what different ‘Surzhyks’ have in common in order to classify them as belonging to a wider, overarching umbrella term of Surzhyk. There is a large variety of what is considered to be Surzhyk in the wider literature, and many authors seek to point to commonalities between different Surzhyk varieties in order to classify them as part of one broader form of Surzhyk. Although the role of either standard Russian or standard Ukrainian is not downplayed in the same way that lexifier languages are downplayed according to Monogenetic approaches, the clear emphasis on commonality between Surzhyks by different

authors writing on the subject demonstrates that Monogenetic approaches can be used in this analysis in order to better understand the wider literature on Surzhyk. One could arguably apply the Nautical-Jargon Theory to the analysis of Surzhyk, classifying it as a language variety with many different regional variations that still have particular linguistic traits in common.

Implications

Using Monogenetic approaches to explain the wide variety of literature on Surzhyk confirms that it can best be seen as a contact language. While Surzhyk cannot be categorized as a pidgin, theories of pidgin development nonetheless help make sense of the various literature on Surzhyk, confirming the categorization of Surzhyk as a contact language. In addition, applying this approach to the analysis of Surzhyk prompts the question of whether Surzhyk can even be considered to have a lexifier language. One could arguably point to standard Russian as Surzhyk's lexifier language, identifying Russian as the source language from which Surzhyk derives much of its lexicon. The question of whether the terminology of a lexifier language applies to the case of Surzhyk will be discussed further in later paragraphs.

Surzhyk and the Creole Continuum

The Concept of a Creole Continuum

Another concept that is useful to apply to this analysis of Surzhyk is the idea of a creole continuum. The theory of a creole continuum is used within contact language research to explain the development of creole languages. The model of the creole continuum seeks to account for the variation found in creole speech production and views creoles as being on a scale relative to their lexifier languages. According to this scale, the basilect is the variety furthest away from the lexifier language, the mesolect is the intermediate stage, and the acrolect is the variety closest to

the lexifier language (Bakker and Matras 2013, 76). An assumption inherent in this concept is that the basilect represents the ‘original’ creole while all further developments along this scale represent a process of decreolization towards the acrolect (Ibid). The creole continuum is but one theory within contact language research that seeks to explain variations in creole development. While some authors view it as a useful model with which to view creole language development, others view this theory as a simplified explanation of creole development, arguing that this continuum does not account for the development of ‘basilectalization’, or inverse development from an acrolect towards a basilect (Bakker and Matras 2013, 76).

Applying the Creole Continuum to Surzhyk

However, the concept of a creole continuum is a useful one when analyzing Surzhyk as a case study. Although it may be a simplified explanation of creole variation and development, this concept of a creole continuum can be a useful theory in explaining and classifying Surzhyk as a linguistic variety. Arguably, Surzhyk can be viewed as a variety that exists somewhere on this continuum as a language variety not unlike a ‘mesolect’, or an intermediate stage between standard Ukrainian and the development of a completely new linguistic variety. As will be discussed later, Surzhyk does not appear to be a creole language, as the concept of a mixed language better describes its structure. However, Surzhyk does appear to share some similarities with the concept of a mesolect, as it is an intermediate language variety that is neither standard Ukrainian nor a fully developed new language. Therefore, using the classification of Surzhyk as a fused lect or developing mixed language one could view Surzhyk as an intermediate stage in the development of a mixed language, not unlike the concept of a mesolect as an intermediate stage in the development of a creole.

Implications

Although the concept of a creole continuum is contested, Surzhyk does share some similarities with the concept of a mesolect, as it is arguably an intermediate language variety that has developed differently from its source languages while still having a degree of mutual intelligibility with standard Russian and Ukrainian. As will be discussed in later paragraphs, Surzhyk lacks the same complexity and established structure as a creole, but by applying the creole continuum to Surzhyk, one can view Surzhyk as an intermediate language variety. These similarities between mesolects and Surzhyk do not indicate that it is a creole, but it remains a useful way of analyzing Surzhyk as a developing contact language. Instead, Surzhyk should be viewed as a developing mixed language that has evolved beyond the stage of a linguistic ‘jargon’ of sorts, with its own set of rules and conventions. However, most mixed languages contain less variation than Surzhyk, meaning that a better, more concrete classification of Surzhyk would be that of a fused lect, or intermediate stage. Comparing Surzhyk to a mesolect of sorts strengthens the classification of this variety as a fused lect, as it demonstrates that Surzhyk is an intermediate stage that has a particular codified system but still contains greater variation than most mixed languages. This classification of Surzhyk as a fused lect or developing mixed language will be discussed further in later paragraphs.

Surzhyk’s Lexifier Language

The Concept of a Lexifier Language

Using the concept of a creole continuum, while recognizing that Surzhyk itself is not a creole language, and applying it to Surzhyk, one could also argue that standard Russian is Surzhyk’s lexifier language. As discussed previously, pidgins and creoles generally derive most

of their lexicon from one language, which is called the lexifier language (Kaye and Tosco 2001, 49). However, mixed languages do not appear to have one important lexifier language in the same way that creoles do. Bakker and Matras differentiate mixed languages from pidgins and creoles by arguing that many mixed languages borrow their lexicon from both source languages equally (Bakker and Matras 2013, 180), although as will be discussed in later paragraphs, this distinction does not apply to intertwined mixed languages that follow a grammar versus lexicon divide and combine the grammar from one source language with the lexicon from another source language.

Applying The Concept of a Lexifier Language to Surzhyk

Applying this concept of a lexifier language to this case study of Surzhyk, several conclusions can be made about the structure and development of Surzhyk. Creoles generally have one main lexifier language from which they acquire the most of their lexicon. A parallel can be drawn between other creole languages and Surzhyk by studying the role of a lexifier language. Many authors argue that Ukrainian provides the base for most of Surzhyk's structure, while much of its vocabulary appears to have been borrowed primarily from standard Russian (Кузнецова, Савченко, and Хмелевский 2018, 95). According to Kent, Surzhyk is a Ukrainian-based variety containing a Russian lexicon and mostly Ukrainian morphosyntax (Kent 2010, 39). If one accepts this assertion of Surzhyk as a primarily Ukrainian language variety with a Russian influence, Russian could be characterized as the lexifier language of Surzhyk, as it provides most of the lexicon. However, as will be discussed further in later paragraphs, Surzhyk can be described better as a mixed language that follows a grammar versus lexicon divide, borrowing fairly equally from both standard Ukrainian and standard Russian. Russian could be viewed as Surzhyk's lexifier language, as it borrows much of its vocabulary from Russian, but as

a language variety comparable to a fused lect or developing mixed language, this label of a lexifier language does not fully describe the position of the Russian language in relation to Surzhyk.

Implications

The concept of a lexifier language is a useful one to apply to Surzhyk, and by framing Surzhyk in these terms, one can view Surzhyk as sharing similarities with creole languages, namely that it is a more developed variety that takes much of its lexicon from one language. However, while the concept of a lexifier language may explain the role of the Russian language in the construction of Surzhyk, Surzhyk's vocabulary does borrow significantly from the standard Ukrainian vocabulary (Flier 1998, 114). Therefore, instead of viewing Surzhyk as a creole with Russian as its lexifier language, it can better be described as a fused lect which takes much of its lexicon from Russian but whose structure is based primarily on Ukrainian. While there is a strong Russian influence on Surzhyk vocabulary, the label of a lexifier language is perhaps not the most applicable to this case, as Surzhyk's lexicon is not 'Russian' to the same extent that the vocabulary of most creoles is based on that of their lexifier language. It would perhaps be more accurate, therefore, to label Surzhyk as an intermediate or developing mixed language.

Mixed Language Development

General Theories of Mixed Language Development

In addition to these theories on pidgin and creole development, there are several theories among researchers of contact languages that attempt to explain the development of mixed languages. Generally, there are two major approaches that attempt to explain the formation of

mixed languages which are useful for this analysis. The first approach is the unidirectional approach, which argues that mixed languages develop as a result of a one-way shift from an ancestor, or source, language to an introduced, or target, language (Viveka 2015, 81). The second approach is the fusional approach, which argues that two source languages combine to form a new third language, thereby creating what is referred to as a mixed language (Viveka 2015, 81). Similar to Bakker and Matras's scale of mixed language autonomy and stability, both the unidirectional and fusional approaches consider large-scale community bilingualism to be a necessary condition for the development of a mixed language (Viveka 2015, 81). According to this theory, unidirectional approaches are one-way theories that describe a language shift from the source language to a target language or from a target language back to the source language. In both cases, the unidirectional approach argues that this shift is not completed, which results in the formation of a mixed language (Bakker and Matras 2013, 194-195). Fusional approaches, however, argue that the two languages merge to create a new mixed language (Bakker and Matras 2013, 194-195), as bilingual speakers of two languages combine elements of these languages in a systematic way (O'Shannessy 2021, 460). Unlike unidirectional approaches, which argue that speakers were in the process of shifting from one language to another and stopped before this process was complete, the fusional approach considers a mixed language to be the result of a combination of two linguistic sources (Viveka 2015, 83).

Applying Theories of Mixed Language Development to Surzhyk

In the context of mixed language theories, Surzhyk can best be classified as a fusional mixed language, or a language variety resulting from the merging of standard Russian and standard Ukrainian to form a new type of mixed language, namely Surzhyk. The definition of a unidirectional mixed language does not apply to Surzhyk. While it could be argued that Surzhyk

developed as a result of a language shift from Ukrainian to Russian as a result of Russification policies in the Soviet Union, Surzhyk can better be viewed as a language variety that developed by combining different elements of Ukrainian and Russian, as explained below. Although Surzhyk could perhaps be explained as a unidirectional mixed language, a better category seems to be that of a fusional mixed language. Like fusional mixed languages, Surzhyk appears to have combined elements of Russian and Ukrainian equally, mixing linguistic standards of both languages (Sira, Di Nunzio, and Nosilia 2019, 5). Surzhyk is a mix of Russian and Ukrainian standards, meaning that elements of both languages are an important characteristic of the nature of Surzhyk (Sira, Di Nunzio, and Nosilia 2019, 5). It involves breaking the linguistic rules of both Russian and Ukrainian (Bernsand 2001, 44) by individuals who are at least somewhat proficient in both languages and therefore combine linguistic elements of Russian and Ukrainian. For example, one common feature of Surzhyk is the combination of Russian prefixes with Ukrainian verbs, as in the frequent use of the prefix под- rather than the Ukrainian equivalent під- in Surzhyk speech (Sira, Di Nunzio, and Nosilia 2019, 2). This mix of a common Russian prefix with many Ukrainian verbs demonstrates that Surzhyk is the result of a combination of two linguistic sources rather than a language shift from Ukrainian to Russian. In addition, the development of Surzhyk predates the Russification policies of the Soviet Union, as Russian has long been a language of status and power in the region. One can find examples of Surzhyk in nineteenth century literature and theater, in which it was used as a comedic representation of rural Ukrainians (Flier 2008, 41). Ukrainian-Russian language mixing is not simply the result of a unidirectional shift to Russian but rather the result of the combination of Ukrainian and Russian linguistic forms by speakers of both languages. Surzhyk can best be viewed as a mix of

standard Ukrainian and Russian, meaning that it shares considerable similarities with fusional mixed languages.

Implications

Like fusional mixed languages, Surzhyk speech combines linguistic elements of both Ukrainian and Russian. Therefore, it appears that Surzhyk has more in common with mixed languages than other contact languages. The category of pidgin languages does not seem to fit Surzhyk as a language variety, and because creoles are argued to have developed from pidgins, this category also seems less applicable to this case. Mixed language development theories, however, appear more applicable to the analysis of Surzhyk. Authors of contact language literature outline a distinction between fusional and unidirectional mixed language development which seems applicable to the case study of Surzhyk. Using this classification, Surzhyk can be explained as a fusional mixed language, as it appears to be a language variety that combines elements and linguistic standards of both Ukrainian and Russian in order to form a new type of mixed language.

Complexity of Surzhyk

Contact Language Complexity

In addition to theories of contact language development, it is useful to discuss classification criteria that are used in defining contact languages as pidgins, creoles, or mixed languages. One classification marker that is often used when analyzing contact languages and classifying them as either a pidgin, creole, or mixed language is that of language complexity. The complexity of a language variety is an important criteria discussed in research about contact languages, meaning that it is a useful distinction to apply to the analysis of Surzhyk. Generally,

pidgin languages are considered to be less complex compared to other contact languages, such as creoles and mixed languages. Pidgins are often characterized as ‘simplified’ or ‘reduced’ languages that have a limited communicative range (Kaye and Tosco 2001, 22). Creoles, however, are seen as more complex languages that have developed from pidgin languages, in the process acquiring more complex language structures. Similarly, researchers of contact languages argue that mixed languages are generally just as complex as their standard language components (Kaye and Tosco 2001, 22). According to Kaye and Tosco, mixed languages gradually become identity markers within a community of speakers, a process that is accompanied by a rise in structural complexity (Ibid), meaning that mixed languages gradually become more complex as they develop, while pidgins remain relatively limited in scope and complexity. Language complexity, therefore, is a useful way of defining and categorizing contact languages. It is generally accepted by researchers of contact languages that pidgins generally cannot express the full range of human communication and are used in very particular contexts, whereas creoles and mixed languages tend to be much more complex varieties that are spoken in a range of different contexts. The complexity of language structure can clarify and explain different possibilities available to speakers of a particular language variety, and as such it is a useful classification criteria in linguistic analysis. Therefore, this criteria will be used in the linguistic analysis of Surzhyk. Linguistic theories on the complexity of pidgins, creoles, and mixed languages can be applied to the analysis of Surzhyk in order to determine some tentative findings about this particular language variety.

Applying Theories of Complexity to Surzhyk

When applying the research on the complexity of contact languages to this particular case study, one can conclude that Surzhyk is arguably a relatively complex contact language when

compared to pidgins, creoles, and mixed languages. However, Surzhyk is also characterized by authors like Del Gaudio as a ‘non-standard ‘language’ whose usage is basically limited to context specific communication spheres” (Del Gaudio 2010, 14). According to this argument, Surzhyk is relatively limited in its communicative scope (Del Gaudio 2010, 73) and lacks the linguistic complexity of standard languages, such as Ukrainian and Russian. Del Gaudio argues that Surzhyk is a compromise between the two standards of these languages, stating that Surzhyk remains limited in its communicative scope (Del Gaudio 2010, 100). However, while Surzhyk is arguably a more simplified language variety than standard Russian or Ukrainian, it still has a degree of complexity comparable to that of a mixed language. One model that has been used specifically in the research on Surzhyk to address its complexity is the MLF model proposed by Sira, Di Nunzio, and Nosilia. The MLF model, as described by these authors, is a framework that analyzes bilingual contact language mixes and describes these linguistic varieties as containing a Matrix Language and an Embedded Language. According to this model, one source language is the Matrix language, which is the source of the morphosyntactic frame in a bilingual clause, while the other source language is the Embedded Language, which adheres to the structural requirements of the Matrix language. This model suggests that Ukrainian is Surzhyk’s Matrix language whereas Russian is an Embedded language (Sira, Di Nunzio, and Nosilia 2019, 2), implying that the structure of Surzhyk is primarily based on standard Ukrainian. Using this classification of Ukrainian as Surzhyk’s Matrix Language, Surzhyk can be viewed as a complex mix that adds a Russian element to its Ukrainian base. Surzhyk’s pronunciation and the basic structure of Surzhyk grammar rely heavily on the source language of Ukrainian (Del Gaudio 2010, 237), but there is a clear influx of Russian influence, in particular in the use of adverbs and Russian-influenced vocabulary in Surzhyk speech (Del Gaudio 2010, 79). If we use this theory

of Ukrainian as Surzhyk's Matrix Language and Russian as its Embedded Language, then we can draw greater parallels between Surzhyk and the category of mixed languages, as they have comparable levels of complexity. Using this theory, Surzhyk is a complex mix of Russian and Ukrainian, meaning that we can conclude that it has greater complexity than other contact languages, such as pidgin languages. One example of this complexity is the use of new linguistic structures in Surzhyk speech. Like other types of mixed languages, in addition to incorporating elements directly from its source languages, Surzhyk also exhibits new linguistic forms that are not directly taken from either Russian or Ukrainian. For example, Surzhyk's morphology mixes elements from Russian and Ukrainian morphology to create new vocabulary. One example of this is the interrogative adverb *why*, which is *чому* in standard Ukrainian and *почему* in standard Russian. In many varieties of Surzhyk, however, the new adverb *чозо* is used (Del Gaudio 2010, 85). This creation of linguistic features in Surzhyk, such as new vocabulary, demonstrates that this variety is not simply a form of code-switching or borrowing of Russian terms into standard Ukrainian speech but rather a complex language variety that has its own unique linguistic forms. Contrary to the opinions held by many Russian and Ukrainian authors writing on the subject, Surzhyk is not a simplistic borrowing of Russian vocabulary in standard Ukrainian speech by monolingual Ukrainian speakers but a complex linguistic variety in its own right (Sira, Di Nunzio, Nosilia 2019, 1).

Implications

Therefore, Surzhyk appears to be a context-specific language variety that lacks the complexity of standard Ukrainian and Russian while still remaining a complex language variety. Using the criteria of language complexity and applying it to this case, it appears that while the use of Surzhyk is perhaps limited to certain contexts or situations, it remains a complex and

varied language variety. Pidgins are generally categorized as far more simplified forms of communication, and although Surzhyk may be less complex than standard Russian and Ukrainian, its functions and the contexts in which it is spoken are more varied than in a pidgin language. Using language complexity as a marker, the label of a pidgin language does not properly define Surzhyk as a linguistic category. Instead, it can be concluded that the characteristics and classification criteria of other contact languages, such as mixed languages and creoles, are more applicable to the analysis of Surzhyk conducted in this thesis.

Stability and Uniformity of Surzhyk

Contact Language Stability

Another criteria that is often used as a marker of language classification in the study of contact languages is the stability and uniformity of a language variety. The classification criteria of stability is one that is often used in the literature in order to place contact languages in one of three categories, defining them as either a mixed language, a pidgin, or a creole, according to the stability of their linguistic structures. Both mixed languages and creoles are generally considered to have a high degree of linguistic uniformity and stability, making them more established and formalized language varieties as opposed to other types of contact languages (Viveka 2015, 82). Pidgin languages, on the other hand, are considered to be more fluid, containing a higher degree of speaker variation. According to the research on contact languages, this increased speaker variation and lack of uniformity makes pidgins less stable language varieties than other contact languages (Kaye and Tosco 2001, 39). Because they lack this individual speaker variation present in pidgins, mixed languages and creoles are considered far more stable as linguistic varieties than pidgins are. Because this criteria is so widely used in the broader research on

contact languages, theories of stability in contact languages offer valuable theories and insights that can be applied to the analysis of Surzhyk. However, it should be kept in mind that there is a high degree of variety in any language and that a lack of standard conventions or grammatical norms among speakers is not unique to contact languages. On the contrary, it can be observed in many ‘standard’ languages as well. Language stability is a criteria that is usually only applied to non-standard contact languages rather than ‘normal’ languages (Bakker and Matras 2013, 200), making the criteria of language stability a flawed criteria. Like standard languages, contact languages are constantly developing and changing over time, and there can be a high degree of variation among speakers of standard languages as well. Mixed languages in particular go through constant linguistic change, as they are often spoken alongside the standard languages in the community, meaning that stability is a difficult criteria by which to judge any contact language (Bakker and Matras 2013, 200). Additionally, the use of stability as a classification marker is a vague criteria in and of itself. In analyzing contact languages, the idea of ‘stability’ is often meant in a relative sense, namely how established linguistic patterns are in a mixed language in comparison to ‘normal languages’ (Bakker and Matras 2013, 200). By comparing the stability of contact languages to ‘normal’ languages, however, researchers of these languages often ignore the changing nature of standard languages and seek to use these languages as a rigid baseline in order to draw a conclusion about the stability of contact languages. Even within the study of contact languages, there are no benchmarks or measures of stability with which to gauge the stability of pidgins, creoles, and mixed languages apart from comparing them to ‘normal’ languages (Bakker and Matras 2013, 200), which is not the most reliable method of analysis. Therefore, although this criteria can offer some insights when applied to the analysis of Surzhyk,

it should be remembered that language stability is an imperfect method of classifying contact languages and should not be the only criteria applied to the analysis of Surzhyk.

Applying Theories of Stability to Surzhyk

Not only is stability a somewhat flawed classification criteria of contact languages in general, relatively little has been written about the linguistic stability of Surzhyk. This is partly due to the degree of variation found in works discussing Surzhyk. Because each author appears to define Surzhyk differently, there is no comprehensive study of how stable or standardized Surzhyk is as a linguistic variety. This makes it difficult to draw concrete conclusions about the stability of Surzhyk, given that further practical research of Surzhyk in Ukraine is difficult to carry out as a result of the military conflict with Russia. However, despite the vague nature of language stability as a classification criteria, it remains a key factor used in the classification and analysis of contact languages. Therefore, the criteria of language stability can be useful in analyzing under-researched language varieties such as Surzhyk, as it allows for further clarification of how Surzhyk may be developing in similar ways to other contact languages. According to the literature on contact languages, there are generally three ways of judging linguistic stability which can be applied to this analysis of Surzhyk. In their discussion of mixed languages, Bakker and Matras outline a scale of linguistic autonomy among contact languages, arguing that most mixed languages fall into one of these three categories, namely complete separation from the source languages, contact but non-fluency in the source languages, and ongoing bilingualism (Bakker and Matras 2013, 201). According to this scale, increased distance from the source language generally promotes greater stability in the structure of a mixed language (Bakker and Matras 2013, 201). However, it should be noted that this is not always the case, especially in situations where speakers of mixed languages are fluent in one or both of the

source languages in addition to speaking a mixed language (Bakker and Matras 2013, 201).

Another measure of stability that is often applied to mixed languages is the degree of consistency between speakers of a particular mixed language variety in their use of lexicon and grammar (Bakker and Matras 2013, 201). If we apply this criteria to Surzhyk, we can argue that there is a relative degree of consistency between different regional ‘Surzhyks’. There may be a great deal of regional variety found in Surzhyk (Del Gaudio 2010, 72), but there are some common elements across these different varieties that lend stability to this language variety. For example, different regional varieties of Surzhyk often exhibit similar pronunciation patterns. One such pronunciation feature unique to most types of Surzhyk is the reduced presence of the Ukrainian *ɛ*, or *v*, at the beginning of a syllable or a word (Del Gaudio 2010, 69). For example, the Ukrainian word *ɛyxо* is pronounced as *(ɛ)-yxо* in Surzhyk, in which the letter *ɛ* at the beginning of the word is much softer than in standard Ukrainian pronunciation (Ibid). Another example of a feature different ‘Surzhyks’ have in common is the “general absence of lengthening in consonantal phonemes in all those phonological contexts where Ukrainian would show a doubling of consonants” (Del Gaudio 2010, 70). Additionally, in Surzhyk speech consonants followed by *я* generally are pronounced as in Russian, although this may differ slightly regionally as well as among individual Surzhyk speakers (Ibid). These examples demonstrate that while there is perhaps a greater deal of regional or individual variety in Surzhyk than in other contact languages, Surzhyk does possess a considerable degree of linguistic stability (Del Gaudio 2010, 230), similar to mixed languages. If we use this scale of mixed language stability and apply it to the analysis of Surzhyk, it can be argued that the best descriptor for Surzhyk is a language variety that falls in the ‘ongoing bilingualism’ category. Similar to many mixed languages, Surzhyk arguably has little distance from its source languages (Del Gaudio 2010,

237). Bakker and Mous's categorization of ongoing bilingualism describes these mixed languages as symbiotic mixed languages, or languages that are spoken alongside the standard languages in a community (Bakker and Mous 1994, 7), meaning that these mixed languages gradually gain stability and autonomy from the source languages while still containing considerable linguistic variation. This description would appear the most applicable of the three categories when analyzing Surzhyk, as Surzhyk is often described in the existing literature as a variety that is spoken alongside standard Russian and standard Ukrainian. Surzhyk is not a language variety that is completely separate from the source languages or a variety that has contact with the source languages but whose speakers are not fluent in those languages. Rather, communities of Surzhyk speakers are arguably bilingual in at least one of the source languages and use either Ukrainian or Russian, if not both languages, frequently in daily life (Sériot 2005, 38). Therefore, one can categorize Surzhyk as a widely spoken contact language which continues to co-exist alongside its source languages, making it a language variety that exists in a situation of ongoing bilingualism in the community.

In addition to their scale describing the linguistic autonomy of mixed languages, Bakker and Matras further describe three indicators of stability in mixed languages, namely independence from their source languages, speaker consistency, and language acquisition (Bakker and Matras 2013, 203). These three indicators of mixed languages are useful for the purposes of the analysis conducted in this thesis, as applying them to this case study can provide a sense of Surzhyk's linguistic stability. However, while the first two indicators are clearly applicable to the case of Surzhyk, there is no substantial literature on the subject of the language acquisition of Surzhyk. Therefore, the indicator of language acquisition is less relevant to this analysis. If one applies the first two indicators to this analysis, then it can be concluded that

while Surzhyk lacks considerable independence from its source languages, namely Ukrainian and Russian, there is a degree of consistency among different speakers of Surzhyk (Del Gaudio 2010, 69-72). When compared to other types of mixed languages, such as Bilingual Navajo, Surzhyk appears to be less stable. Whereas these mixed languages are considered to have relatively stabilized structures that are common among its different communities of speakers, Surzhyk arguably contains more variation. For example, in the case of Bilingual Navajo, because the ancestral language provides the essential structure and the introduced language provides the lexicon (Mazzoli and Sippola 2021, 3), there is a clearer set of accepted standards and norms used by most speakers of this mixed language. In Surzhyk, however, this clear delineation between the contribution of the ancestral and introduced languages is not present, and the structure and lexicon may vary somewhat in different Surzhyk-speaking communities. However, as discussed by Del Gaudio, Surzhyk does appear to have particular norms and language structures that are found in different regional examples of Surzhyk (Del Gaudio 2010, 69-72). Language acquisition of Surzhyk may differ according to the regions in which it is spoken, but one can conclude that Surzhyk speakers do have particular language norms in common, such as patterns of pronunciation.

Implications

Using language stability as a classification factor in the analysis of Surzhyk, it appears that Surzhyk has more in common with mixed languages than with pidgins or creoles. Therefore, it seems a near definitive conclusion that we cannot classify Surzhyk as a pidgin. Pidgins lack consistency and stability across speaker communities, and contain far greater linguistic variation. As described by the literature on contact languages, pidgins are an intermediary linguistic phase used for the purposes of context-specific communication, and as such are far more fluid than the

other contact languages discussed in this thesis, lacking formalized structure and speaker conventions. Surzhyk, however, seems to have more in common with mixed languages and perhaps creoles, as it has a greater degree of established norms and standards across different regional varieties, which is something that pidgin languages lack. This places Surzhyk in the category of other contact languages such as mixed languages and creoles, meaning that theories about these language varieties are the most useful in analyzing Surzhyk's stability and uniformity as a linguistic variety. Bakker and Matras's indicators of mixed language stability, therefore, appear to be the most applicable for the purposes of this research, providing greater insight on Surzhyk's stability and uniformity as a contact language. Using these theories, Surzhyk appears to be a fairly stabilized and consistent language variety with both a fair degree of regional variety and a number of common characteristics across different 'Surzhyks'. This puts it into the category of contact languages such as creoles and mixed languages. This tentative classification of Surzhyk will be further discussed in later paragraphs.

Mixed Language Categories

Intertwined and Converted Mixed Languages

In addition to the application of different classification criteria to this analysis of Surzhyk, it is also useful to discuss different categories of mixed languages in order to better offer a tentative classification of Surzhyk as a mixed language. Mixed languages are generally placed into several categories. One such category is the distinction between intertwined languages and converted languages. Intertwined languages are languages that are made up of two mutually dependent components (Viveka 2015, 71). Each of these components originates from a different source language, meaning that intertwined mixed languages rely on both of these

components to create linguistic meaning. Usually, these two components of intertwined languages follow a grammar versus lexicon divide, meaning that most of their grammar originates from one of the source languages while most of the lexicon originates from the other source language (Viveka 2015, 71). Converted languages, however, are languages that have adopted the grammar of another language without changing the lexicon, thereby adapting the grammar of one language to the lexicon of another (Viveka 2015, 75). As a result, converted mixed languages have a different structure than intertwined mixed languages. For example, while Michif, an intertwined mixed language, combines the verbs and verbal system of Cree with the nouns and nominal system of French (Viveka 2015, 73), converted mixed languages such as Sri Lanka Portuguese adopt the grammar of one language, namely Tamil, while keeping the lexicon of Portuguese (Bakker 2000, 33). In addition to this distinction between converted and intertwined languages, other authors make a distinction between classical types, split types, split reverse types, and reverse types of mixed languages (Bakker and Matras 2013, 213). Classical types of mixed languages are languages in which the morphosyntax is from the original community language and the lexicon is from the new or introduced language (Mazzoli and Sippola 2021, 29). Split types of mixed languages are generally considered to be languages that take grammatical structures from both source languages, such as borrowing the verbal system from the ‘original language’ and the nouns from the ‘introduced language’ (Ibid). Split reverse types of mixed languages, however, are the opposite, in which the nouns are from the original community language while the verbs are generally borrowed from the ‘introduced language’ (Ibid). Finally, reverse types of mixed languages are those that have a grammar-lexicon divide in which the lexicon is from the original community language and the morphosyntax is introduced to form a new language (Ibid).

Surzhyk as an Intertwined Classical Mixed Language

Using these theories, we can define Surzhyk as a type of intertwined mixed language, or a language that combines mutually dependent elements of both source languages, in addition to a fusional mixed language. Using these categories to analyze Surzhyk, one could view Ukrainian and Russian as ‘mutually dependent’ languages in the creation of this language variety. As discussed in the previous paragraph, components of both languages are combined equally to form the linguistic structures of Surzhyk. Additionally, applying other categories of mixed languages to the analysis of Surzhyk, one could view Surzhyk as a classical type of mixed language. Like many classical mixed languages, Surzhyk arguably follows a grammar versus lexicon divide, as much of the grammatical structure of Surzhyk is taken from Ukrainian while a great deal of the lexicon is borrowed from Russian (Del Gaudio 2010, 237). Therefore, one could view Surzhyk as a classical type of mixed language, as indicated by the views of Del Gaudio and others. In the definition of Surzhyk referred to here, the primary syntax is arguably Ukrainian (Flier 1998, 114), so much so that some authors view it as a Russified dialect of Ukrainian rather than a true contact language (Del Gaudio 2010, 237). Surzhyk arguably combines Russian vocabulary with Ukrainian grammar and pronunciation (Кузнецова, Савченко, and Хмелевский 2018, 95). Although the lexicon of Surzhyk also borrows from the lexicon of standard Ukrainian, it demonstrates the most Russian influence, as the Surzhyk lexicon borrows most substantially from Russian vocabulary. As a result, it can be observed that Surzhyk most resembles an intertwined and classical mixed language. Like other intertwined mixed languages, Surzhyk appears to be made up of two mutually dependent components rather than completely adapting the grammar of one language to the lexicon of another, as the grammar and lexicon of Surzhyk still demonstrate a combination of both Ukrainian and Russian elements.

Implications

As a result, applying the categories of mixed languages outlined in the literature to this case study of Surzhyk, the most applicable contact language classification is that of a mixed language. In particular, the category of intertwined classical mixed languages appears the most applicable to Surzhyk. Like these types of mixed languages, Surzhyk appears to be made up of two mutually dependent languages which both contribute equally to the structure of Surzhyk along a grammar and lexicon divide, in which the primary structure of Surzhyk grammar is taken from Ukrainian and Surzhyk's primary lexicon is taken from Russian. As a result, when compared to other contact language categories, such as those of pidgin and creole languages, the classification of a mixed language seems the most applicable to Surzhyk.

Symbiotic Mixed Languages

Plain and Symbiotic Mixed Languages

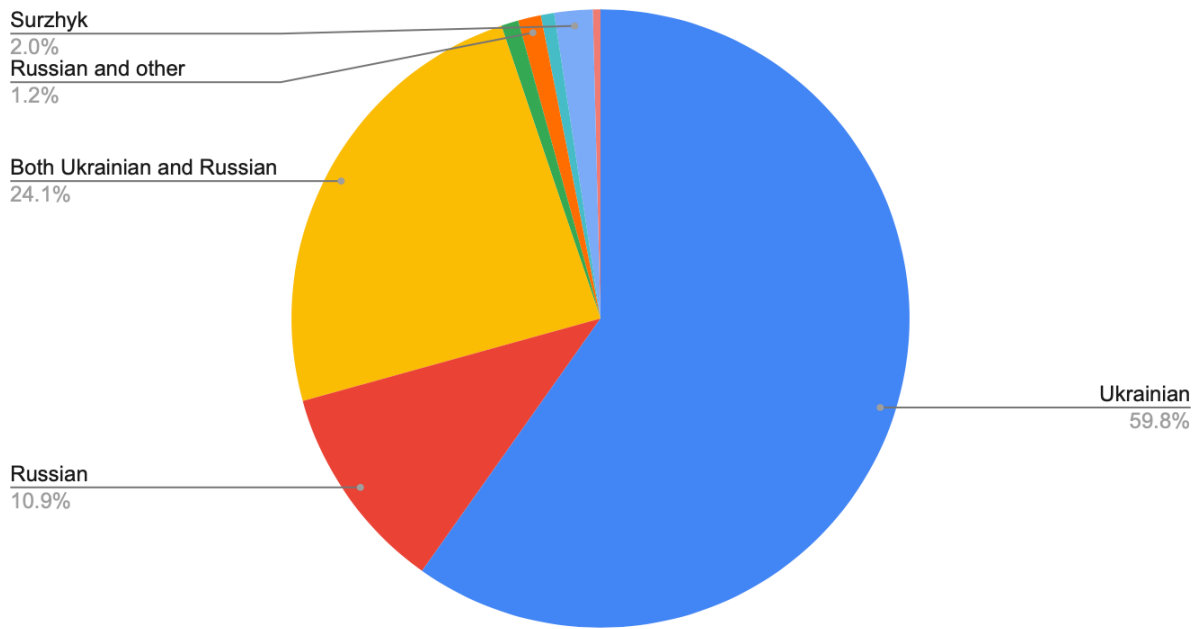
In addition to these categories of mixed languages, another useful category of mixed languages that can be applied to the case study of Surzhyk is that of symbiotic mixed languages. Bakker and Mous differentiate between plain mixed languages and symbiotic mixed languages. According to this definition, plain mixed languages are generally the only language spoken in a particular community, while symbiotic mixed languages are spoken alongside the dominant source languages (Bakker and Mous 1994, 7). As a result, symbiotic mixed languages are also considered to have a more limited lexicon than plain mixed languages (Ibid). Symbiotic mixed languages, therefore, are mixed languages that are spoken alongside the source languages in the community, meaning that most speakers in the community are effectively bilingual in both the dominant source languages and the mixed language. This is not the case for all mixed languages,

but, as discussed previously, most mixed languages are generally spoken in a situation of ongoing community bilingualism. It is rare for a mixed language to be spoken in total isolation from its source languages, and as a result, most mixed language speakers are fluent in one or both of the source languages (Bakker and Matras 2013, 200-201). As a result, most mixed languages can be classified as symbiotic mixed languages which coexist alongside the dominant languages in a particular speech community, making plain mixed languages the minority.

Surzhyk as a Symbiotic Mixed Language

Surzhyk has many characteristics in common with mixed languages, and it shares many similarities with symbiotic mixed languages in particular. Although there are some monolingual speakers of Surzhyk (Bilaniuk 2004, 422), like other symbiotic mixed languages, Surzhyk is generally spoken in the context of community bilingualism. According to Sériot, most Surzhyk speakers are trilingual, speaking both Ukrainian and Russian in addition to Surzhyk (Sériot 2005, 38). According to the graph below, the majority of Ukrainians report Ukrainian as their native language, while a smaller but still significant percentage report that both Ukrainian and Russian are their native languages. As a result, it can be inferred that most Surzhyk speakers are fluent or at least proficient in both source languages.

Figure 1: Self-Reported Native Language in Ukraine



Ukrainian: 59.8%
Russian: 10.9%
Both Ukrainian and Russian: 24.1%
Ukrainian and other: 0.9%
Russian and other: 1.2%
Other: 0.7%
Surzhyk: 2.0%
Difficult to answer: 0.4%

Data derived from Bureiko and Moga 2019, 144

Using Bakker and Mous's distinction of symbiotic versus plain mixed languages, one can argue that Surzhyk can be tentatively classified as a symbiotic mixed language. Symbiotic mixed languages in particular usually develop in situations where a common language already exists and in which communication is not an issue (Bakker and Mous 1994, 9-10). This can easily be compared to Surzhyk, as speakers of Surzhyk generally also have access to either Russian or Ukrainian and speak one or both of these languages fluently or at least proficiently. While it is theoretically possible to observe monolingual speakers of Surzhyk, this is seen primarily in older populations in rural areas of Ukraine (Del Gaudio 2010, 257). Like other symbiotic mixed languages, the use of Surzhyk did not develop in response to a communication gap but rather for a specific social purpose in a bilingual community (Del Gaudio 2010, 237). The general consensus among researchers of Surzhyk is that it is generally spoken alongside Russian or Ukrainian and that most Surzhyk speakers are proficient in one if not both of these source languages (Bilaniuk 2004, 422), meaning that the category of symbiotic mixed languages seems the most applicable to the analysis of Surzhyk.

Implications

Therefore, it seems clear that Surzhyk can best be classified as a mixed language. While theories of pidgin and creole development may be useful in the analysis of this case study, Surzhyk shares the most similarities with mixed languages. The different categories of mixed language appear best to explain Surzhyk as a nonstandard language variety. While the definition of Surzhyk as a mixed language is an imperfect definition, it is the most useful one when assigning Surzhyk a particular label as a contact language.

Surzhyk as an Identity Marker

Contact Languages as Identity Markers

An additional feature of contact language classification that is commonly used among researchers of contact languages is the question of identity and how this relates to a particular language variety. One common feature of mixed languages in particular is that they are considered to be an identity marker by linguists. Most researchers of contact languages agree that mixed languages appear to have been created consciously as an act of identity (Mazzoli and Sippola 2021, 47), rather than to solve a language barrier, as is the case with pidgin and creole languages (Bakker and Matras 2013, 181). As a result, mixed languages are often an expression of a particular community identity, either a completely new identity or a different interpretation of an older identity (Bakker and Matras 2013, 181). Therefore, mixed languages are often spoken not for the purposes of straightforward communication but for the purpose of in-group identity. As discussed in the previous paragraph, most mixed languages are spoken alongside other languages in the community. As a result, most speakers of mixed languages do not rely on these language varieties as their primary language of communication but instead use it to signal a particular group identity.

Surzhyk as a Form of In-Group Identity

This feature of mixed languages can easily be applied to Surzhyk. While Surzhyk can arguably be a speaker's primary language, this is rare and more often Surzhyk appears to be spoken primarily as an identity marker. Surzhyk can be viewed as a stigmatized variety that comprises part of an important local and rural identity for many Ukrainians (Yekelchik 2010, 224). As seen in the graph above, two percent of Ukrainians report Surzhyk as their native language. However, the number of speakers for whom Surzhyk is a native language may be

underreported because of the stigmatization of Surzhyk. Many Ukrainians may report either standard Ukrainian or Russian as their native language rather than Surzhyk because they may not want to admit to speaking a variety of low status. In addition, the question of language nativity is highly subjective and relies on an individual's complex understanding of their linguistic identity. Surzhyk speakers may perceive one of the standard languages as their primary language while they in fact use Surzhyk in most of their communication (Bilaniuk 1997, 108), or they may not wish others to know that their native language is Surzhyk. As a result, monolingual speakers of Surzhyk may be underrepresented in this graph. However, it is clear that Surzhyk is definitely a form of identity, much like other mixed languages. Surzhyk is arguably used less for communication purposes but rather as a form of identity within a specific in-group. It is a specific part of linguistic and cultural identity, meaning that an important feature of Surzhyk is its function as an identity marker. We can use the case of Verka Serduchka as an example of how Surzhyk is used within Ukraine as a specific identity marker. Verka Serduchka is the stage name of Andriy Mykhailovych Danylko, who became famous as a cross-dressing comedian and performer. Surzhyk in the media is often associated with Serduchka (Del Gaudio 2010, 277), and her songs are an example of how Surzhyk is used within Ukraine to construct a particular form of identity. In Serduchka's songs, "the performance of gender, ethnicity, and language is highly stylized and is fraught with ambiguities: a female character is played by a male performer; the performer seamlessly draws on Russian, Ukrainian, and Surzhyk" (Tovares 2019, 468). Serduchka used Surzhyk as part of a particular identity she constructed as a performer. By using Surzhyk as an identity marker to signal a particular form of Ukrainian rural identity, Serduchka created a fluid image for herself as a performer (Ibid). She used Surzhyk to construct this rural peasant persona, signaling its role as an identity marker to speakers of this variety. However,

Serduchka's use of Surzhyk for the purposes of comedy and entertainment might be seen as mocking, and later in Serduchka's career, she no longer used Surzhyk speech as part of her persona, although this also may have been due to the stigma surrounding the use of this language variety. "Ever since Verka the lowly train attendant became Verka the pop star, the proportion of Surzhyk in her utterances diminished considerably" (Yekelchuk 2010, 224). In addition, the Surzhyk used most often in Serduchka's songs was not 'authentic' Surzhyk following the particular norms and conventions of Surzhyk but rather a type of code-switching aimed to establish the character's identity (Yekelchuk 2010, 224). It is difficult to know exactly why Serduchka no longer uses Surzhyk as a part of this entertainment persona, but the fact that she stopped doing so seems significant, demonstrating that it is a particular identity marker, albeit perhaps a stigmatized one that was viewed negatively by a wider Ukrainian audience. However, Serduchka's use of Surzhyk in her older performances illustrates Surzhyk's role as a part of a specific in-group identity by showcasing this language variety as part of Serduchka's constructed rural identity. Surzhyk, like other mixed languages, is therefore demonstrably part of a community identity.

Implications

Using theories of identity formation among mixed language communities, therefore, is helpful in the analysis and classification of Surzhyk. Unlike creoles and pidgins, mixed language theory focuses much more on their status as identity markers among speaker communities. Mixed languages in general are very much expressions of a particular community identity, something that is applicable to this case study of Surzhyk. Surzhyk is also used primarily as an expression of a particular stigmatized rural identity in Ukraine. Therefore, as discussed in previous paragraphs, this confirms the classification of Surzhyk as a type of mixed language.

Surzhyk as a Native Language

Contact Languages and Nativity

Another criteria that is used in the classification of contact languages is whether a particular language variety can be considered to be a native language. Language nativity is an indicator of the contexts in which particular contact languages are spoken, giving an indication of how they can best be classified. Generally, creoles and mixed languages can be considered native languages, at least to some extent, while pidgin languages are considered by researchers to be varieties that cannot be spoken as native languages. According to most researchers, creole is a pidgin which has acquired native speakers (Kaye and Tosco 2001, 34). According to this definition, a pidgin is a non-native contact language with a far more limited lexicon and more limited communicative functions, while creoles are pidgins that have developed as native languages of particular communities. Mixed languages are slightly more difficult to categorize according to this criteria, and there are differing opinions among researchers on whether they should be considered as native languages for their group of speakers. According to Bakker and Matras, the majority of mixed languages are the native language of a community (Bakker and Matras 2013, 186), even if they are often symbiotic mixed languages that are spoken alongside their source languages. Language nativity is also often seen by contact language researchers as an indicator of language stability. Although mixed languages can arguably be native languages, many mixed languages have only elderly speakers and no child learners (Bakker and Matras 2013, 202), meaning that the criteria of language nativity is not always applicable to mixed languages. Mixed languages may also be native languages at one point but then lose this status over time as a result of linguistic stigma or standardization.

Can Surzhyk Be a Native Language?

Using this criteria of contact language classification and applying it to the case study of Surzhyk can give us an indication of where Surzhyk falls on the contact language spectrum. There is much disagreement on the question of whether Surzhyk can and should be considered a native language. As seen in the graph above, a small number of Ukrainians report Surzhyk as their native language, while the majority report either Ukrainian or both Ukrainian and Russian as their native language. As discussed previously, however, the number of native or monolingual speakers of Surzhyk may in fact be much higher, as Surzhyk's low status prevents many speakers from claiming it as their native language. The question of whether Surzhyk can be spoken as a native language is a complex one, in part because there is a great deal of variation between different 'Surzhyks'. According to most of the research on this linguistic variety, however, Surzhyk can be a native language, depending on the context. There are Ukrainians for whom Surzhyk is their primary language, making it arguably a 'native' language of sorts for these speakers. According to Sériot, older generations can be native speakers of Surzhyk, while younger generations use Surzhyk more rarely in informal contexts (Sériot 2005, 45). There are Surzhyk speakers who are not fluent in either Russian or Ukrainian, meaning that for these speakers, Surzhyk as a fused lect is arguably their native language (Bilaniuk 2004, 422). These speakers of Surzhyk could be described as absolute Surzhyk speakers, according to Del Gaudio's distinction between non-absolute, absolute, and semi-absolute Surzhyk speakers. According to this definition, the absolute Surzhyk speaker is a person whose exclusive language of communication is Surzhyk (Del Gaudio 2010, 257). These speakers are mainly of older generations and tend to live in country villages (Del Gaudio 2020, 182). This is rare, however, as in most cases Surzhyk appears to be a symbiotic mixed language that is spoken alongside Russian or Ukrainian by speakers who are at least somewhat proficient in these standards. There

are also speakers of Surzhyk who speak one of or both standard languages but for whom the use of Surzhyk is the norm of their speech community (Bilaniuk 2004, 422). Del Gaudio defines these speakers as semi-absolute Surzhyk speakers, who predominantly use Surzhyk but who are able to switch to either standard Russian or standard Ukrainian with a certain degree of proficiency (Del Gaudio 2010, 257). In contrast, a non-absolute Surzhyk speaker is someone “who might have a good command of either Russian or Ukrainian, or sometimes of both languages, reverting to Surzhyk only under specific social and psychological situations, or else as a consequence of mental associations” (Del Gaudio 2010, 257). These ‘non-absolute’ Surzhyk speakers are generally of a relatively well-educated younger generation (Del Gaudio 2010, 257). Using these definitions, Del Gaudio posits that native Surzhyk speakers are people whose first language was Surzhyk, and who had little exposure to either Russian or Ukrainian during their childhood (Del Gaudio 2010, 258). Using Del Gaudio’s definitions of absolute, non-absolute, and semi-absolute Surzhyk speakers, Surzhyk can be spoken as a native language in particular contexts, especially among older rural Ukrainian generations. While younger speakers of Surzhyk for the most part do have a command of one of the source languages, Surzhyk is a primary language of communication for many speakers. Surzhyk can be a first language for some speakers as the language of their particular community, even if they speak it alongside the Russian or Ukrainian standard, meaning that it is arguably a native language for many Surzhyk speakers.

Implications

Using this classification of language nativity, Surzhyk appears similar to contact languages such as mixed languages and creoles. Because pidgins are generally considered to be varieties in an intermediate stage of language development that are created for specific

communication purposes, they cannot be a speaker's native language. Creoles and mixed languages, by contrast, can be viewed as native languages in the varying literature. Surzhyk, therefore, fits more into these language categories. As argued above, we can better see Surzhyk as a mixed language than as a creole. This is because while the differing literature considers creoles to be a speaker's native language, mixed languages can both be non-native and native languages, depending on the particular mixed language and the contexts in which it is spoken. Just as mixed languages can be native languages for some and non-native languages for others, depending on the context in which the language developed and by whom it is spoken, Surzhyk can be a native language for semi-absolute or absolute Surzhyk speakers and a non-native language for non-absolute Surzhyk speakers.

Surzhyk as a Codified Linguistic System

Mixed Language Theory and 'Fused Lects'

Using mixed language theory, one can describe Surzhyk as a codified linguistic system with particular rules. Mixed languages are not a form of random linguistic mixing but in fact autonomous codified linguistic systems with particular rules, although these systems may be fluid and changing. In contact language literature, fused lects are generally viewed as an intermediate step in the development of a mixed language (Mazzoli and Sippola 2021, 95). Mazzoli and Sippola posit that code-switching becomes 'code-mixing' in a community of bilingual speakers, before developing into a 'fused lect', followed by the development of an independent mixed language (Mazzoli and Sippola 2021, 95).

Surzhyk as a ‘Fused Lect’

According to this definition of ‘fused lects’, it could be argued that Surzhyk should be classified not only as a mixed language but more specifically as a developing mixed language, or a fused lect. Contrary to the opinions of many Russian and Ukrainian authors, Surzhyk is not a form of incorrect Russian or Ukrainian but rather a codified system. Like other mixed languages and ‘fused lects’, Surzhyk’s linguistic system follows an internal logic. The structure of Surzhyk is neither arbitrary nor artificial (Flier 1998, 130) but a fused lect that follows an “inner logic of mixing and corrupting words” (Yekelchuk 2010 224). Surzhyk is not simply random jargon or mixed speech; just as mixed languages are autonomous linguistic systems with particular rules, Surzhyk also follows particular rules and conventions within the community of its speakers. While these rules are somewhat fluid, as Surzhyk may develop and change in the same way that standard ‘normal’ languages continually change, it remains a codified linguistic system similar to a fused lect or developing mixed language.

Implications

While mixed languages tend to exhibit a high level of uniformity in their lexicon and structure (Viveka 2015, 70) Surzhyk tends to contain more regional variation. In this way, Surzhyk deviates somewhat from the ‘traditional’ conception of a mixed language, as it contains more variation than most mixed languages. For example, as discussed in previous paragraphs, mixed languages such as Bilingual Navajo have a clear set of accepted standards and norms used by most speakers of this mixed language while Surzhyk has more regional variation. Therefore, the more specific definition of Surzhyk as a fused lect or developing mixed language appears more relevant to the analysis of Surzhyk. Surzhyk contains more variation than most mixed languages, but it has many similarities with mixed languages, meaning that a better classification

of Surzhyk might be that of a fused lect. Therefore, Surzhyk can best be considered as a fused lect, or developing mixed language, which follows an internal logic and has a codified structure while containing more regional variation than other, more developed, mixed languages.

Conclusion

Classification of Surzhyk

The aim of this thesis has been to offer a classification of Surzhyk as a language variety. The existing literature on Surzhyk offers many different definitions and classifications of Surzhyk. There is no single definition of this phenomenon that is accepted by different authors, meaning that the term *Surzhyk* is used to refer to a number of different language phenomena in Ukraine. Because of this, there are a number of different classifications of Surzhyk which operate from a different understanding of Surzhyk. However, although it can be argued that each author uses the term *Surzhyk* differently, these definitions overlap with one another. It can be said that there are different ‘Surzhyks’ which are described differently in the literature, but there are enough similarities between these different definitions of Surzhyk to apply them to this research. As such, by applying the academic theories of contact language research to the analysis of Surzhyk, there are several conclusions that can be drawn. Firstly, Surzhyk can best be described as a contact language. Although some authors, such as Del Gaudio, view Surzhyk as a form of a Ukrainian dialect or sociolect rather than a true contact language, the classification of Surzhyk as a contact language remains useful for the purposes of this thesis, as contact language research remains the most applicable in attempting to explain Surzhyk’s particular linguistic nature. Surzhyk shares some similarities with pidgins and creoles as well as mixed languages, and many of the classification criteria that are applied to these contact languages can be applied to Surzhyk.

While these categories of pidgin, creole, and mixed languages are but overarching definitions that overlap to a certain extent, meaning that they are somewhat fluid, classification criteria of contact languages offer useful insights to the classification of Surzhyk. Although Surzhyk can better be defined as a fused lect or developing mixed language than a creole, it does share some similarities with creole languages. As a result, it can be concluded that Surzhyk shares some similarities with pidgin and creole languages as well as mixed languages. Applying the concept of a creole continuum to Surzhyk, we can conclude that as a fused lect, Surzhyk shares some similarities with a mesolect, or an intermediate creole language that is on a continuum between standard Ukrainian and a new contact language. Using this comparison to a creole language, we can tentatively describe Russian as Surzhyk's lexifier language, the language that provides the base of Surzhyk's vocabulary, although this definition does not fully encompass the variety present in Surzhyk's vocabulary. However, the category of a mixed language more accurately describes the unique characteristics of Surzhyk. While pidgins are described as less complex and stable language varieties, Surzhyk is arguably fairly complex and has a certain degree of stable linguistic norms that are shared by its different regional varieties. Surzhyk can be described as a fusional, intertwined, and symbiotic mixed language, as it is a language variety that appears to be composed of two mutually dependent components, following a grammar versus lexicon divide, that is spoken alongside the source languages in the community. Although these classifications do not perfectly describe Surzhyk, they are perhaps the most useful in explaining Surzhyk's unique linguistic features. Like other mixed languages, Surzhyk can have native speakers, but it is more often spoken alongside its source languages in the community as a form of an in-group identity. Surzhyk is not a form of random mixing of languages but rather a codified system that follows particular rules, similar to a fused lect. As a result, while it shares some similarities with

pidgin and creole languages, Surzhyk can best be classified as a fused lect or developing mixed language. This classification is by no means a concrete definition, but rather a tentative classification that may explain some of Surzhyk's idiosyncrasies.

Future Research

The scope of this research has been limited by recent events. The Russian invasion of Ukraine has impacted the ability to conduct field research on Surzhyk, and it has further changed the linguistic landscape of Ukraine. It is evident that the use of the Russian language, as well as the use of Surzhyk, is now viewed more negatively as a result of Russia's position as the invader. As a result of this shifting view of the Russian language, the social and linguistic contexts of Surzhyk are likely to have changed. Surzhyk has already been stigmatized as an 'impure' or Russianized form of Ukrainian, and this is likely to have increased as a result of the war with Russia. Therefore, field research on the subject of Surzhyk is needed. It has been outside the scope of this research to investigate the sociolinguistic effects of the war with Russia on Surzhyk, but future research would benefit from investigating this topic. Because of the recent changing linguistic landscape in Ukraine, the sociolinguistic contexts in which Surzhyk is spoken may have also changed. The linguistic structure of Surzhyk may also have changed, and more research is needed to investigate this. There has been a greater emphasis on the use of Ukrainian instead of Russian, and further research should investigate the effect of an increasing move to the Ukrainian language on the structure and use of Surzhyk. Another interesting topic that should be further investigated is the use of Surzhyk among Ukrainian refugee communities. As a result of the conflict in Ukraine, these refugee communities have become established in Europe, and their language use may demonstrate an emphasis on the increased use of standard Ukrainian rather than Russian or a non-standard language variety such as Surzhyk. Surzhyk is

already a stigmatized variety within Ukraine, and this has likely increased as a result of the war with Russia. Therefore, further research is needed in order to investigate the ongoing effects of the Russian-Ukrainian conflict on the structure of Surzhyk and its use within Ukrainian refugee communities. These topics are outside of the scope of this thesis, but future research of Surzhyk should address these issues.

Bibliography

- Bakker, Peter and Mous, Maarten. 1994. *Mixed Languages: 15 Case Studies in Language Intertwining*. Amsterdam: IFOTT.
- Bakker, Peter. 2000. "Convergence Intertwining: An Alternative Way towards the Genesis of Mixed Languages." In *Languages in Contact*, 29-35. Amsterdam: Rodopi.
- Bakker, Peter and Matras, Yaron. 2003. *The Mixed Language Debate: Theoretical and Empirical Advances*. Berlin/New York: de Gruyter Mouton.
- Bakker, Peter and Matras, Yaron. 2013. *Contact Languages: A comprehensive guide*. Berlin: de Gruyter Mouton.
- Bernsand, Niklas. 2001. "Surzhyk and National Identity in Ukrainian Nationalist Language Ideology." *Berliner Osteuropa Info* 17: 38-47.
- Bernsand, Niklas. 2006. "The othering of Surzhyk in Ukrainian implicit metalinguistic discourse." *History, Language and Society in the Borderlands of Europe* 1: 77-115.
- Bilaniuk, Laada. 1997. "Speaking of "Surzhyk": Ideologies and Mixed Languages." *Harvard Ukrainian Studies* 21, no. ½: 93-117.
- Bilaniuk, Laada. 2004. "A typology of Surzhyk: Mixed Ukrainian-Russian language." *The international journal of bilingualism* 8, no. 4: 409-425.
- Bilaniuk, Laada. 2018. "Purism and Pluralism: Language Use Trends in Popular Culture in Ukraine since Independence." *Harvard Ukrainian Studies* 35, no. ¼: 293-309.
- Bureiko, Nadiia and Moga, Teodor Lucian. 2019. "The Ukrainian-Russian Linguistic Dyad and its Impact on National Identity in Ukraine." *Europe-Asia Studies* 71: 137-155.
- Del Gaudio, Salvatore. 2010. *On the Nature of Suržyk: a Double Perspective*. München: Sagner.

- Del Gaudio, Salvatore. 2015. "Украинско-Русская Смешанная Речь "Суржик" в Системе Взаимодействия Украинского и Русского Языков." *Полемика* 2: 214-246.
- Del Gaudio, Salvatore. 2020. "Language situation in the district of Ripky (Černihiv, Ukraine)." *Russian Linguistics* 44, no. 2: 177-201.
- Долешаль, У., Дубичникий, В.В., and Ройтер, Т. 2011. "Суржик: Лексико-Грамматический Анализ." *Русский язык в научном освещении* 2: 247-267.
- Flier, Michael S. 1998. "Surzhyk: The Rules of Engagement." *Harvard Ukrainian Studies* 22: 113-136.
- Flier, Michael S. 2008. "Suržyk or suržyks?" In *Belarusian Trasjanka and Ukrainian Suržyk: Structural and social aspects of their description and categorization*, 39-60. Oldenburg : Bibliotheks- und Informationssystem der Universität Oldenburg.
- Fomina, Joanna. 2014. "Language, Identity, Politics - The Myth of Two Ukraines." *Institute of Public Affairs*: 1-20.
- Gasparov, Boris. 2006. "Suržyk: A Glance From a Personal Perspective." *Harvard Ukrainian Studies* 28: 117-128.
- Groen, Ben. 2014. "The Language Question in Contemporary Ukraine." In *Definitely Perfect*, 185-222. Amsterdam: Uitgeverij Pegasus.
- Kaye, Alan S. and Tosco, Mauro. 2001. *Pidgin and Creole languages: A basic introduction*. München: LINCOM Europa.
- Kent, Kateryna. 2010. "Language Contact: Morphosyntactic Analysis of Surzhyk Spoken in Central Ukraine." *Working Papers in Linguistics* 8: 33-53.
- Курохтина, Т.Н. 2009. "Суржик в Интернете." *Вестник Московского Университета* 9: 101-112.

- Кузнецова, И. В., Савченко, А. В., and Хмелевский, М. С. 2018. “Киевский Суржик как Феномен Украинской Городской Разговорной Речи (На Примере Комедии М.П. Старицкого за Двумя Зайцами»)." *Вестник ЧГПУ* 2: 92-102.
- Lakhtikova, Anastasia. 2017. “Understanding Passive Bilingualism in Eastern Ukraine.” *Critical Multilingualism Studies* 5: 144-173.
- Mazzoli, Maria and Sippola, Eeva. 2021. *New Perspectives on Mixed Languages: From Core to Fringe*. Berlin/Boston: de Gruyter Mouton.
- McConvell, Patrick. 2008. “Mixed Languages as Outcomes of Code-Switching: Recent Examples from Australia and Their Implications.” *Journal of language contact* 2: 187-212.
- Nazarenko, Lilia. 2012. “Украинский Суржик - Социолект, Просторечие, Пиджин? К Вопросу о Языковой Ситуации на Украине.” *Slavica Litteraria* 15: 243-250.
- O’Shannessy, Carmel, Hakimov, Nikolay, and Auer, Peter. 2021. “How ordinary child language acquisition processes can lead to the unusual outcome of a mixed language.” *The international journal of bilingualism* 25, no. 2: 458-480.
- Оснач, Сергій. 2019. “Кому варто боятися закону про мову?” *Портал мовної політики*, May 17, 2019.
<http://language-policy.info/2019/05/komu-var-to-boyatysya-zakonu-pro-movu/>
- Romanova, Ekaterina, Zhironkina, Oksana and Vakhtin, Nikolai. 2007. “Suržyk: In Search of a Research Object.” *Harvard Ukrainian Studies* 29: 327-348.
- Sériot, Patrick. 2005. “Diglossie, bilinguisme ou mélange de langues: les cas du suržyk en Ukraine.” *La Linguistique* 41, no. 2: 37-52.
- Shevchuk, A. O. 2021. “Russianisms (Surzhyk) As One of the Main Problems of Media Texts.” *POLIT. Challenges of science today*: 359-360.

- Sira, Nataliya, Di Nunzio, Giorgio Maria, and Nosilia, Viviana. 2019. "Towards an Automatic Recognition of Mixed Languages in R: The Case of Ukrainian-Russian Hybrid Language Surzhyk." diss., University of Padua.
- Solomon, Erika. 2022. "For Russian-Speaking Ukrainians, Language Clubs Offer Way to Defy Invaders." *New York Times*, May 29, 2022.
<https://www.nytimes.com/2022/05/29/world/europe/ukraine-russia-language.html?referrerSource=articleShare>
- Stadnik-Holzer, Elena. 2007. "Epistolary Surzhyk: A Contribution to the Study of a Ukrainian Substandard." *Harvard Ukrainian Studies* 29: 349-369.
- Taranenko, Oleksandr. 2007. "Ukrainian and Russian in Contact: Attraction and Estrangement." *International Journal of the Sociology of Language* 183: 119-140.
- Tovares, Alla V. 2019. "Negotiating "thick" identities through "light" practices: YouTube metalinguistic comments about language in Ukraine." *Multilingua* 38, no. 4: 459-484.
- Viveka, Velupillai. 2015. *Pidgins, Creoles, and Mixed Languages: An Introduction*. Amsterdam: John Benjamins Publishing Company.
- Yekelchuk, Serhy. 2010. "What Is Ukrainian about Ukraine's Pop Culture?: The Strange Case of Verka Serduchka." *Canadian-American Slavic studies* 44: 217-232.
- Зинкевич, Андрей. 2013. "Языковые контакты в Славии – XXI-й век." *Wiener Slavistisches Jahrbuch* 1: 129-149.