https://doi.org/10.15407/ukrmova2022.02.038
UDC 811.161.1'27:811.161.2'27

JAN PATRICK ZELLER, Dr. Professor at the Chair for Slavic Linguistics,
University of Greifswald, Germany
Ernst-Lohmeyer-Platz 3, 17487 Greifswald
Tel.: +49 (0)3834 420 3222
E-mail:jan.zeller@uni-greifswald.de
https://orcid.org/0000-0002-6942-7112

OKANNIA AND *AKANNIA* IN UKRAINIAN-RUSSIAN MIXED SPEECH ("SURZHYK")

The Ukrainian-Russian mixed speech (URMS), also known as "surzhyk", is a widespread phenomenon in central areas of Ukraine. Linguistic studies still lack empirical research on the variation of phonic characteristics of URMS and on its connection with the social characteristics of its speakers. Based on a corpus of spoken speech of this non-standard variety with around 340,000-word tokens taken from informal family conversations and open interviews, this article examines the variation in Ukrainian-Russian mixed speech between two prominent phonic features of Ukrainian and Russian — the variation of unstressed /ɔ/ between Ukrainian *okannia* and Russian *akannia*. The results confirm that *okannia* strongly dominates in unstressed vocalism of URMS thus largely corresponding to the standard Ukrainian pattern and differentiating it from Russian.

A Generalized Linear Mixed Model shows that the variation between *okannia* and *akannia* is influenced by complex dialectal and sociolinguistic differences in the Ukrainian language landscape, sociodemographic characteristics of the speakers, e.g., age and gender, and the speech situation. There are evident correlations between the phonic variation and the lexical-morphological affinity of the word form, i.e., whether the word form on the lexical-morphological level coincides with either standard Ukrainian or Russian. The findings make it clear that one-dimensional attempts to clarify language variation in Ukraine are bound to fail.

Keywords: Ukrainian-Russian language contact, dialect contact, language variation, okannia, akannia

1. INTRODUCTION

The Ukrainian-Russian mixed speech (URMS), commonly derogatorily referred to as "surzhyk", is a widespread phenomenon in Ukraine, especially in its central areas. Most observers assume that its phonic (phonetic-phonological) side is strongly based on (dialectal) Ukrainian (Del Gaudio, 2010; Taranenko, 2013); however, we find variation between Ukrainian and Russian

C i t e s: Zeller, J. P. (2022). *Okannia* and *akannia* in Ukrainian-Russian mixed speech ("surzhyk"). *Ukraïns'ka Mova, 2*(82), 38—59. https://doi.org/10.15407/ukrmova2022.01.038

features both between speakers and within the speech of one and the same speaker. There is still a lack of empirical studies on the variation of phonic characteristics in URMS and on its connection with the sociodemographic characteristics of its speakers (like age, gender, education). Thus, the present study represents a first step in this direction. Based on a corpus of spoken speech, the present study examines the variation of one of the most striking phonic differences between Ukrainian and Russian: the pronunciation of unstressed /ɔ/ after non-palatalized consonants — i.e., the variation between *okannia* and *akannia*.

In the further course of the introduction, I briefly present the phenomenon of URMS as well as the state of research on phonic variation in URMS and discuss *okannia* and *akannia* in Ukrainian, Russian and some of their subvarieties. In Section 2, I provide information on the corpus that serves as the data basis and explain the methods for the analysis. Section 3 presents the results of a Generalized Linear Mixed Model analysis of the variation between *okannia* and *akannia* in the corpus. In section 4, the results are discussed and summarized.

1.1. Ukrainian-Russian mixed speech

Speech containing elements (representations/realizations of lexical items, wordforms, morphemes, phonemes) and structures that coincide partly with elements/structures of the Ukrainian language and partly with elements/structures of the Russian language is widespread in Ukraine, especially in its central areas (Del Gaudio, 2010; Hentschel & Taranenko, 2015). Such speech is commonly referred to as "surzhyk". Due to the negative connotation of this word, I will instead use the term "Ukrainian-Russian mixed speech". Despite forms of mixed speech having existed for a long time in Ukraine, as a phenomenon that affected large parts of the population, URMS arose in the 20th century, when the influence of standard languages was only fully felt among the masses of the population. Due to the Russian-oriented language policy that existed during most of the Soviet era, many Ukrainians oriented themselves towards the Russian standard language rather than the Ukrainian standard language during most of the 20th century. Another decisive factor in this development was the rapid urbanization in Ukraine after the Second World War, when many formerly rural speakers of Ukrainian dialects entered linguistic surroundings dominated by Russian.

Sometimes, URMS is described as a chaotic mix of Ukrainian and Russian, and attributed to an alleged lack of education and language proficiency of its speakers. However, today, many people speak URMS alongside Ukrainian and Russian (Hentschel & Taranenko, 2015). Quite often these are well-educated children of the original dialect speakers, who are, in fact, able to speak at least one of the standard languages competently. These speakers thus do not use URMS due to a poor proficiency in the standard languages; for them, the URMS is rather an informal, family variety in their linguistic repertoire.

Hentschel (2014, 2018) has emphasized that URMS is a part of a European-wide trend. In the 20th century, convergence between autochthonous dialects and standard languages could be observed in many European linguistic communities (Auer, 2005). The result of these processes are mesolectal variants (regiolects and urban dialects), which are situated between old, small-scale autochthonous dialects acting as basilects and the standard language acting as an acrolect [1]. Such mesolectal varieties can become quite stabilized, in the sense of a reduction in their variability, resulting from the "eradication of socially or locally marked variants" (Watt & Milroy, 1999, p. 26), and they can become focused varieties (cf. Le Page & Tabouret-Keller, 1985), i.e., an awareness can arise in the speech community of what is and what is not a "normal" way of speaking, and thus, a linguistic norm (see also Kerswill, 2007). This still includes speakers potentially varying their speech both towards the acrolect and towards the basilect depending on the situation (Auer, 1986; Schilling-Estes, 2002). In the case of URMS, its only specific feature is that, for most of the 20th century, the standard language that functioned as the acrolect was not the standard language structurally closest to the dialects, i.e., Ukrainian, but rather Russian. Still, the Ukrainian standard language also played a role in Ukraine to a much higher degree than the Belarusian standard language did (and does) in Belarus, where the so-called "trasianka" represents a similar phenomenon to that of Ukrainian "surzhyk". In addition, since Ukraine's independence, standard Ukrainian has slowly but steadily gained prestige and acceptance in the country. One can, therefore, surmise that the convergence to Russian is coming to an end. Furthermore, in the south and east of Ukraine, where the language landscape is shaped by Russian, a "neo-surzhyk" could emerge as a result of speakers turning from Russian to Ukrainian, but such mixed speech would certainly be more spontaneous than the traditional URMS on a Ukrainian dialectal basis in Ukraine's central regions, which has already existed for some generations (Hentschel & Reuther, 2020).

It is known from variationist linguistics (e.g., Labov, 1972; Trudgill, 1986) that linguistic features vary in normal language use and that this variation correlates with the sociodemographic characteristics of the speakers and the characteristics of the speech situation. This also applies in cases referred to as "dialect contact" — i.e., in cases where, to some degree, mutually intelligible and structurally close linguistic varieties meet. In such a situation, linguistic accommodation (i.e., the gradual and temporary adaptation of speech to the interlocutor or other linguistic models) is enabled. In this sense, the language contact between Ukrainian and Russian in Ukraine, with URMS as one of its outcomes, resembles situations that can be described as dialect contact.

[1] These terms are often associated with Bickerton's (1975) analysis of the variation between Guyanese Creole (as the basilect) and English (as the acrolect), with intermediate stages between the two as mesolectal varieties. If I use these terms here (following Auer, 2005) with reference to dialect-standard contact, it is by no means my intention to imply that the phenomenon of URMS (or the Ukrainian dialects) can be likened to a creole language. From a contact linguistic point of view as well as from a linguistic-structural point of view, it is simply incorrect to classify URMS as a form of creole (see Voss, 2008a, 2008b; Del Gaudio, 2010; Hentschel, 2014).

URMS has hardly been examined quantitatively. A database for this kind of quantitative research in terms of a corpus of authentic linguistic material was also lacking. Under the direction of Gerd Hentschel, University of Oldenburg, such a corpus of URMS has been created as part of two research projects [2]. This corpus considers the central regions of Ukraine, where URMS is assumed to be strongest. Hentschel (2018) offered an initial insight into the quantitative relationships in this corpus, concentrating on those utterances that were linguistically "mixed", i.e. contained both elements corresponding to the Ukrainian standard language and elements corresponding to the Russian standard language. He showed that, for the overwhelming majority of speakers, the proportion of words corresponding to the Ukrainian standard language among specific words (i.e., word forms that are not identical in both languages) was between 50 and 80 percent. Thus, URMS is significantly more "Ukrainian" than Belarusian-Russian mixed speech (the aforementioned "trasianka") is "Belarusian" [3]. In Section 2, this corpus, which also provides the basis for the present study, is presented in more detail.

1.2. Phonic variation in URMS

The phonic (phonetic-phonological) side of URMS has also thus far barely been examined quantitatively. For all cases of regular phonic differences between Ukrainian (including its dialects) and Russian, the question of variation between different phonetic realizations arises. Most observers assume that the phonic side of URMS corresponds strongly to the Ukrainian language (Del Gaudio, 2010, p. 65; Trub, 2000, p. 53; Taranenko, 2013, p. 40). This is plausible in view of the history of its origin in a partial language shift by many individuals from dialectal Ukrainian towards Russian. However, there are also other voices. Šumarova (2014, p. 317), for example, noted many phonic influences of Russian in URMS. Yet, this has not been checked on the basis of a large database.

Zeller (2018) presented an initial empirical, quantitative study of the aforementioned URMS corpus. Since his goal was an initial comparison with Belarusian-Russian mixed speech, Zeller limited his study to phonic features (in a broader sense) that are also relevant for Belarusian-Russian language contact (e.g., a fricative where standard Russian has a velar plosive /g/, non-palatalized affricates [ʧ] vs. Russian [ʧʲ], prosthetic sounds before back vowels, [l] vs. an approximant or fricative in certain lexemes and as a suffix in the past tense masculine). It should be noted that not all of the features discussed by Zeller (2018) are purely phonic, i.e., not all of them can be described synchronously without turning to other linguistic levels. On the one hand, the aim of the study was to

---

[2] *Inflectional morphological irregularity/-ies in 'current' contact varieties of North Slavic languages* funded by the Deutsche Forschungsgemeinschaft and *Variability and stability in mixed substandard: Suržyk* funded by the Fritz Thyssen Stiftung. As part of the project *Hybridization from two sides: Ukrainian-Russian and Russian-Ukrainian Code Mixing in the context of the (socio)linguistic situation in southern Ukraine along the Black Sea coast*, led by Gerd Henschel and Tilmann Reuther and funded by the Deutsche Forscgungsgemeinschaft and the Wissenschaftsfonds Österreich, the corpus is currently being expanded to include oblasts on the Black Sea coast.

[3] The proportion of "Belarusian" word forms (i.e., those corresponding to the Belarusian standard language) in Belarusian-Russian mixed speech is between 20 and 60 percent for most speakers (Hentschel, 2018, p. 134).

get a more differentiated and graded picture of the phonic side of URMS by asking which individual phonic variables were less and which were more prone to influence by the Russian language. On the other hand, another goal was to examine whether speakers who used more 'Russian' elements at structurally deeper levels (lexicon, morphology) also did so at the level of pronunciation, and furthermore, whether this pronunciation was connected with the so-called "lexical-morphological affinity" of the word in which the variable appeared [4]. The latter refers to whether the word form — apart from its pronunciation — coincides with the corresponding word form in the Ukrainian and/or Russian language (see below and Hentschel, 2018). It was found that purely phonic variables clearly tended towards a Ukrainian-like pronunciation (U [ɦ] / [ɣ] vs. R [g]; U [ʊ̯] vs. R [v]), regardless of the speaker and the affinity of the word. The exception was the variable (ʧʲ), for which — almost without exception — palatalized realizations were noted in the corpus [5]. A greater degree of variation was found in the lexically/morphologically determined variables (U [ʊ̯] vs. R [l], U [ʊ̯] vs. R [u], U prothetic [v]). For these variables, the phonic influence of Russian was generally rather small, but, depending on the group of speakers (whether or not the speaker tended more towards Russian on the lexical-morphological level) and the affinity of the word form, this influence varied in intensity. Correlations between the realization of the phonic variables and sociodemographic characteristics of the speakers were not the focus of Zeller (2018), and some important phonic characteristics have not yet been examined.

1.3. Unstressed vowels in Ukrainian and Russian: *okannia* and *akannia*

One of the most striking differences between the phonic systems of Ukrainian and Russian is the difference between Ukrainian *okannia* and Russian *akannia*. In Russian, the opposition between unstressed /ɔ/ and unstressed /a/ is neutralized after non-palatalized consonants and in the syllable-initial position due to the merger of their phonetic realizations (Avanesov, 1956, p. 106; Panov, 1979, p. 156; Timberlake, 2004, p. 45; Kasatkin, 2006, p. 167; Yanushevskaya & Bunčić, 2015, p. 225) [6]. This phenomenon is called *akannia* ("saying [a]"; R *akan'e*). The phonetic realization of unstressed /ɔ/ and /a/ depends on the position of the vowel. In the immediately prestressed syllable as well as in the syllable-initial position and, optionally, in the absolute final sound of a phonological word, both /a/ and /ɔ/ are realized as [ɐ] or [ʌ]. In all other unstressed syllables, they are realized as [ə].

[4] Single quotation marks are used to indicate that the classification of the lexical-morphological affinity of words or utterances as 'Ukrainian', 'Russian', 'common' or 'hybrid' is done "from the outside (etic) perspective" using a well-defined algorithm (cf. Hentschel, 2018). The classification does not need to correspond completely to the internal (emic) perspective of the speakers themselves.

[5] The palatalized pronunciation of this affricate seems to be a general development in the Ukrainian language. It is not necessarily due to Russian influence.

[6] After palatalized consonants, the opposition between /ɔ/ and /a/ as well as /ɛ/ and /i/ is neutralized in the unstressed position in Russian as well. All are pronounced as an [i]-like sound (the so-called *ikan'e*).

In contrast, in Ukrainian, /ɔ/ is realized as [ɔ] or [o] in unstressed syllables so that the opposition to /a/ is maintained (Rusanovskii et al., 1986, p. 10 and p. 14; Hryshchenko et al., 2002, p. 53; Pompino-Marschall et al., 2017, p. 6). This phenomenon is called *okannia* (R *okan'e*).

Okannia is also characteristic of northern Russian dialects. The fact that these are under the influence of the Russian standard language provides important comparative material in regard to the Ukrainian situation (see below). Directly relevant to the present study is that, while most Ukrainian dialects behave like the standard language in having *okannia, akannia* can be found in some Ukrainian dialects as well. These are the northern Ukrainian (Polissian) dialects, and particularly the eastern parts of this dialectal area, i.e., the region of Chernihiv, and northern parts of the Sumy region and of the Kyiv region (Zhylko, 1966; Del Gaudio, 2017). This area is often described as a transition zone between the Belarusian dialect zone, which has different types of *akannia*, and the Ukrainian dialect zone (Del Gaudio, 2020).

There are some studies on the contact between the *okannia*-pattern and the *akannia*-pattern in East Slavic languages and varieties. *Okannia* has been described as a typical feature of Russian as spoken in Ukraine (Bondarko & Verbitskaia, 1987; Chertorizhskaia, 1988; Del Gaudio, 2011) and of regional variants of Russian found in Russia (Bondarko & Verbitskaia, 1987; Erofeeva, 1997). An elaborate empirical study on the variation of *okannia* and *akannia* was presented by Kochetov (2006), who studied the influence of social variables in Pokcha, a small town in the Western Urals, on the pronunciation of unstressed /ɔ/. He found that the oldest speakers in his sample (born between 1910 and 1923) almost exclusively showed dialectal *okannia* in their speech, whereas standard-like *akannia*-realizations occurred only rarely. In all, [ɔ]- or [o]-like realizations made up 95% of all tokens in the speech of the oldest age group. For middle-aged informants (born between 1929—1947), *okannia* still dominated, but was slightly less common (85 %). A large difference was then found when looking at the speech of the youngest speaker group (born between 1960—1990). *Okannia* was still present but made up only 40 % of all tokens. Kochetov associated this sound change, which began between the 1930s and 1950s, with the dramatic upheavals taking place in the region during that time: "collectivization" and the "cultural revolution," the Gulag, World War II, and the migration of villagers into town (Kochetov, 2006, p. 115).

2. METHODS AND DATA

2.1. Corpus

The basis for the present study is a corpus of URMS taken from family conversations and interviews. The speakers came from the traditional central areas of Ukraine as well as the adjacent regions of Khmel´nyc´kyj in the west, Dnipropetrovs´k (Dnipro since 2016) in the south-east, and Kharkiv in the east. The Family subcorpus contains recordings of free, informal and sponta-

Table 1. **Overview of the Corpus**

Subcorpus	Word Tokens	Informants (Male/Female)	Tokens per Informant	Towns *	Tokens per Town
Family subcorpus	172,482	85 (39/46) (56 with more than 500 word tokens)	Range: 5—11,000 Median: 997.5 Mean: 2005.6	7	Range: 22,191—27,335 Median: 24,760 Mean: 24,640
Interview subcorpus	169,054	68 (28/40)	Range: 813—7311 Median: 2342 Mean: 2486	24	Range: 1092—18,421 Median: 5606 Mean: 7044

* Two towns appear in both the Family and Interview subcorpus. The total number of towns is 29.

neous conversations among friends and family. The Interview subcorpus contains recordings of open interviews with speakers who, in a previous survey, had indicated using URMS on a regular basis. Table 1 provides an overview of the subcorpora.

It is important to emphasize that not all passages in the corpus are "mixed", i.e., not all include both 'Ukrainian' and 'Russian' elements. The speakers also produced utterances and even longer passages that completely coincided with standard Ukrainian as well as utterances corresponding to Russian, i.e., they switched or shifted between Ukrainian and URMS and, to some degree, Russian. In general, it therefore makes sense to differentiate between "URMS in a narrower sense" and "URMS in a broader sense" (cf. Hentschel & Zeller, 2012, in relation to the Belarusian trasianka). The former would be URMS understood as the realization of a "mixed" linguistic code in the speakers' linguistic repertoire next to Ukrainian and Russian. 'Hybrid' utterances are realizations of this code, but utterances, completely corresponding to standard Ukrainian or Russian, can also be realizations of this code, especially if they are short and contain only a few specific (not 'common' to both languages) linguistic elements. On the other hand, mixed speech can be understood in a broader sense as mixed discourse in which speakers switch or shift between Ukrainian, URMS in the narrower sense, and Russian [7]. In the following, I will not differentiate between 'hybrid', 'Ukrainian', and 'Russian' utterances but take the whole mixed discourse of the informants into account.

2.2. Transcription and annotation

The transcription in the corpus of URMS is a broad one. Regarding unstressed /ɔ/, the transcription distinguishes between <o> ([ɔ], possibly [o]) and <a> ([a], [ɐ], [ʌ], or [ə]). Word stress is also transcribed in the corpus. For each word token, the corpus also contains the corresponding word form in standard Ukrainian and/or Russian.

[7] This is not to be confused with Del Gaudio's (2010) distinction between a prototypical, i.e., Ukrainian-Russian-mixed, "surzhyk" and "surzhyk" as a term for other/all types of mixed or non-standard speech.

For the purpose of this study, the linguistic data was automatically annotated using a script programmed in R version 4.1.1 (R Core Team, 2021). The input for the script was the transcription in the corpus and the annotated corresponding Ukrainian and/or Russian word form. Any <o> that was not marked as stressed and did not occur after a consonant marked as palatalized was scored as an instance of *okannia*. For <a>, the task was somewhat more complex, since "unstressed" <a> (reflecting [a], [ɐ], [ʌ], or [ə]) could correspond not only to U and R /ɔ/, but also to U and R /a/. It was only scored here if it corresponded to /ɔ/. Therefore, each <a > not marked as stressed after a consonant not marked as palatalized was matched with the corresponding vowel in a comparison form: If the word form was classified as 'Ukrainian' or 'common', the comparison form was the corresponding Ukrainian word form. If it was classified as 'Russian' or 'hybrid', the comparison form was the corresponding Russian word form. If an <o> was notated for the corresponding vowel in the comparison form (and not <a>), then the vowel was scored as an instance of *akannia*. The script was based on string operations provided by the stringr package in R (Wickham, 2019).

2.3. Statistics

Statistical analysis was performed by means of Generalized Linear Mixed Models (GLMM, see Baayen, 2008) using the lme4 package (Bates et al., 2015) in R (R Core Team, 2021). This method made it possible to simultaneously model and test the influence of various potentially relevant fixed factors, e.g., age, gender, etc., and, at the same time, consider the influence of so-called random factors — here, above all, the influence of individual speakers. The term "random factor" refers to the fact that the data were not independent of each other, but were uttered by a limited number of speakers. Furthermore, these speakers contributed to the corpus to varying degrees. The GLMM method ensured that incorrect conclusions regarding the relevance of factors such as age or gender were not drawn if differences between population groups in the corpus were only due to the characteristics of individual speakers who were quantitatively strongly represented. Such models are not easy to calculate computationally, especially if they also contain interactions of different fixed factors. Therefore, I controlled only the influence of the individual speaker (informant) and was not able to consider other possible random factors, e.g., family, lexeme, or word token.

The dependent variable was a binary one: *okannia* vs. *akannia*. The models calculated changes in the logits, i.e., logarithm of the probability of *akannia* divided by the probability of *okannia*, attributed to the influence of the explaining fixed factors. Based on these logits, which are not easy to interpret intuitively, one can calculate and compare the general probability of producing an *okannia*-realization or an *akannia*-realization for different types of speakers, i.e. for speaker types with certain sociodemographic characteristics.

2.3.1. Sociodemographic and stylistic factors

In terms of sex or gender, many studies have shown that women and men behave differently when faced with language variation. The influence of age on linguistic variation is also commonly found. The connection between these factors and general language usage in central Ukraine was investigated by Hentschel and Zeller (2017) as well as Zeller, Taranenko, and Hentschel (2019) on the basis of a survey from 2014. The authors found no significant difference between male and female respondents with regard to their most commonly used language/code. Differences in general language usage were found to correlate with age, but the effect was small and did not relate to URMS. Younger informants somewhat less often indicated using Ukrainian in favor of Russian, which can be interpreted as a long-term effect of the Soviet era. The number of respondents who indicated using mainly URMS did not differ between older and younger respondents. Still, this does not exclude the possibility that there are differences and changes between female and male, or younger and older respondents, within their URMS itself.

The difference between the two subcorpora, i.e., the Family and the Interview corpus, can be seen as a stylistic one: The interviews with an unknown interviewer were certainly more formal from the perspective of the participants, and people in formal situations tend to show more standard-like speech. On a global level, Hentschel (2018) did not find any differences between the two subcorpora, contrasting with the situation in Belarus, where people tended to make their Belarusian-Russian mixed speech more Russian-like in interviews. However, he noted that, particularly at the beginning of the interview, the interviewees tended to make their speech more Ukrainian-like.

Importantly, given the regional differences in the Ukrainian language landscape, it was assumed that such factors as age, sex (gender), and subcorpus do not have the same effect everywhere in the surveyed area, but depend on the area from which the speakers originated.

2.3.2. Dialectal/sociolinguistic area

For the sake of analysis, the area under investigation was divided into three parts. This division was partly motivated by the traditional dialects present in these subareas and partly by the contemporary sociolinguistic situation. Firstly, regarding the dialectal division of the area under investigation, the relevant factor is the distinction between the northern subarea, where *akannia* is characteristic of the corresponding dialects, and the rest of the surveyed area, where the dialects are characterized by *okannia*. The northern subarea contained only two locations in the region of Chernihiv (the locations in the Sumy and Kyiv regions present in the corpus do not fall into the northern dialect area characterized by *akannia*). Secondly, the division of the large remaining surveyed area was motivated by current linguistic preferences, as determined by Hentschel and Taranenko (2015) on the basis of the aforementioned survey. According to the respondents' assessment in the survey regarding the frequency of use of the three codes — Ukrainian, Russian, and

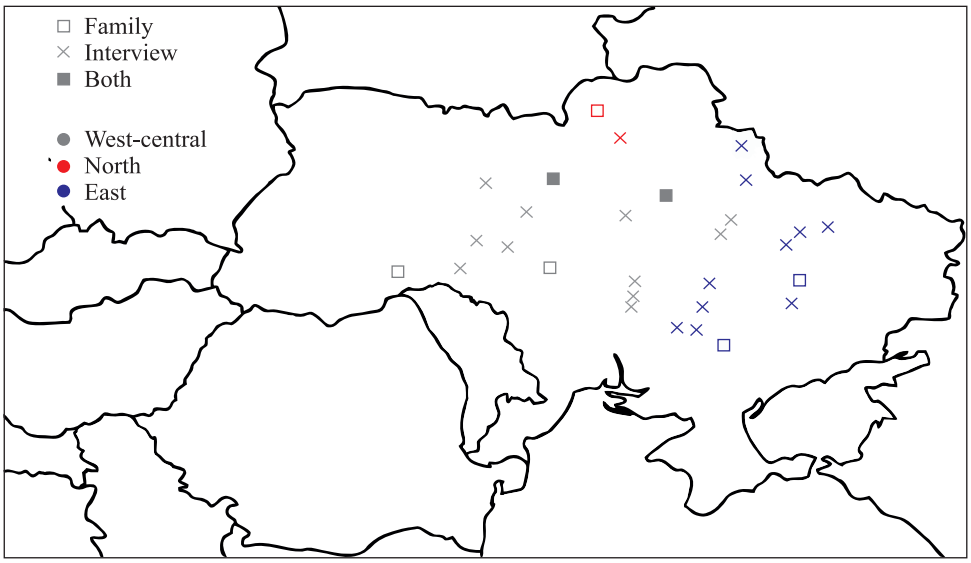

Fig. 1. Survey Locations and Areal Division

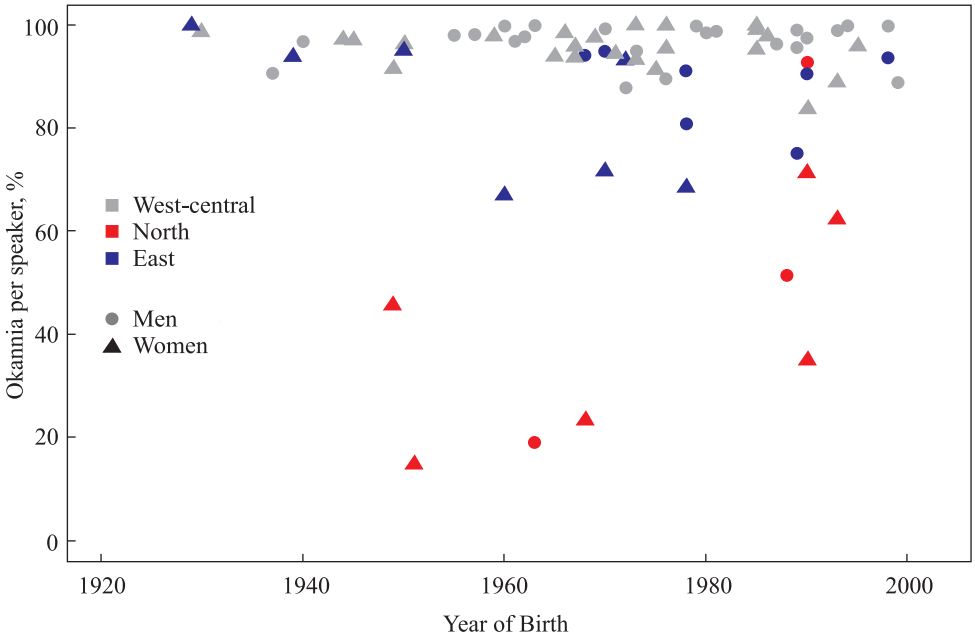

Fig. 2. Individual Variation in the Family Corpus

URMS — the authors divided the surveyed area into five parts. Since there are not that many locations presented in the corpus, their classification had to be simplified here. Therefore, the rest of the surveyed area was divided into a west-central subarea on the one hand and an eastern subarea on the other hand. The west-central subarea contained the regions A, B, and C according to Hentschel and Taranenko (except for their aforementioned northern parts,

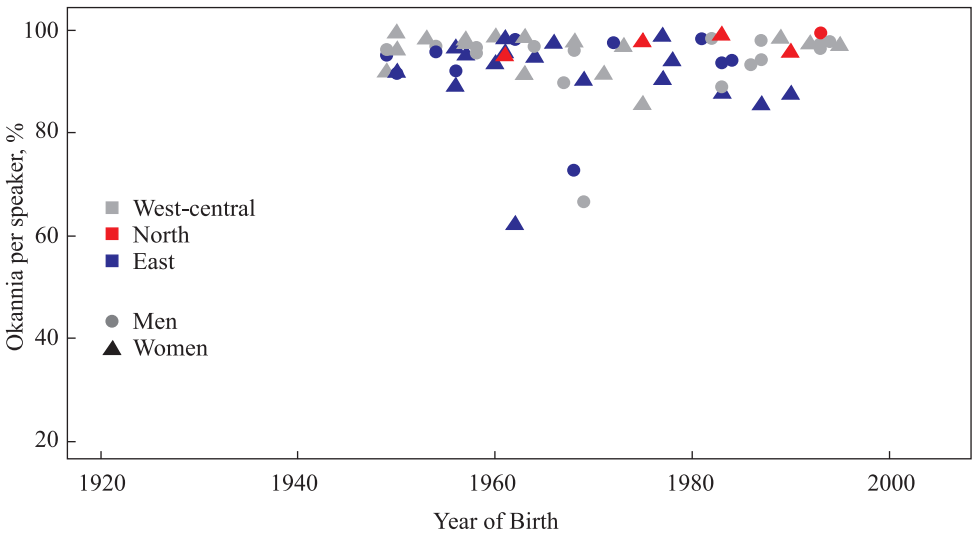

Fig. 3. Individual Variation in the Interview Corpus

where *akannia* can be found). Here, Ukrainian dominates as the mainly used language — albeit to varying degrees. In the eastern area — which encompasses the Kharkiv, Sumy, and Dnipro regions — either the Russian language (Kharkiv; region E according to Hentschel and Taranenko) or URMS (Sumy, Dnipro; region D) dominates. The areal division is shown in Figure 1.

2.3.3. Lexical-morphological affinity

In addition to evaluating the sociodemographic factors, the study also checked to what extent the variation between *okannia* and *akannia* was lexically driven, i.e., related to the correspondence of a word token with the Ukrainian or Russian language. A differentiation was made according to whether the word form in which the instance of the phonic variable occurred — apart from purely phonic characteristics of the word form — corresponded to the Ukrainian and/or Russian language, and, accordingly, was classified as 'Ukrainian', 'Russian', 'hybrid', or 'common' in the corpus (cf. Hentschel, 2018).

3. RESULTS

3.1. Overview

Table 2 gives an initial overview of the data. Note that the number of tokens does not sum up to a total of 80,299 tokens for all factors due to missing values or, in the case of the lexical-morphological affinity, some other, quantitively irrelevant values.

The Ukrainian pattern clearly predominates. In nine out of ten cases in the corpus, an unstressed /ɔ/ is phonetically realized as [ɔ], whereas a pronunciation that corresponds to standard Russian is very rare. The figures in Table 2 hint at only minor influences of sociodemographic factors. The most

pronounced deviation from the general pattern is found in the northern area, where *akannia* is also present dialectally. Also, 'Russian' words have the highest proportion of pronunciations consistent with Russian *akannia*. Nevertheless, for all values of all factors, Ukrainian *okannia* clearly predominates.

Figures 2 and 3 give insight into the individual variation between speakers. Only speakers with at least 20 vowel realizations are depicted. The vast majority of speakers have a clear tendency towards the Ukrainian pronunciation pattern. Again, at first glance a correlation with social factors does not appear to be evident, except for the areal factor in the Family corpus.

3.2. Subcorpus and area

Since it was expected that the effect of factors such as age and gender would depend on the area from which the speakers originated, I first examined whether the influence of areal origin was comparable in both subcorpora. Since this was not the case — as was already indicated in Figures 1 and 2 above and will be shown statistically below — further analyses were carried out separately according to the subcorpus. This also served to simplify the analyses and their interpretation. Otherwise, three-way interactions between subcorpus, area, and the other factors would have had to be tested and interpreted [8].

Table 3 shows the results of a Generalized Linear Mixed Model testing the influence of area and subcorpus on the vowel realization, with informant (n = 152) as a random factor (number of observations: 80,299).

The analysis showed that the influence of areal division on the vowel realization was different in the two subcorpora. As can be seen in Table 3, for the Family corpus (the reference level), a significant difference between the west-central and eastern areas as well as between the west-central and northern areas was found. Releveling the factors to other reference values also revealed that *akannia* realizations were significantly less probable in the east compared to in the north in the Family corpus ($b = -2.03$, $SE = 0.33$, $z = -6.14$, $p < 0.001$). It should be reiterated that family was not treated as a random factor and that there was only one family from the northern area present in the Family corpus. In the Interview corpus, *akannia* realizations turned out to be slightly more probable in the east compared to in the west-central area ($b = 0.60$, $SE = 0.23$, $z = 2.57$, $p = 0.010$) and in the north ($b = 1.05$, $SE = 0.35$, $z = 2.99$, $p = 0.003$). There was no difference between the northern and west-central areas in the Interview corpus ($b = 0.45$, $SE = 0.44$, $z = -1.04$, $p = 0.30$). Also, there was no difference between the subcorpora in the west-central area (which was the reference level in the analysis, see Table 3). The subcorpora differ significantly in the north ($b = -3.87$, $SE = 0.40$, $z = -9.71$, $p < 0.001$) and in the east ($b = -0.79$, $SE = 0.26$, $z = -2.99$, $p = 0.003$), with *akannia*-realizations being less probable in the Interview corpus.

[8] Originally, this was attempted. However, most models did not converge. Therefore, this approach had to be abandoned.

Table 2. **Overview of the Distribution of *Okannia* and *Akannia***

Subgroup	*Okannia* (n)	*Akannia* (n)	Total (n)	*Okannia* %	*Akannia* %
Total	72,113	8186	80,299	89.8	10.2
Subcorpus					
Family	32,557	5406	37,963	85.8	14.2
Interview	39,556	2780	42,336	93.4	6.6
Gender					
Men	29,862	2620	32,482	91.9	8.1
Women	42,251	5566	47,817	88.4	11.6
Age in 2014					
<35	18,203	2331	20,534	88.6	11.4
<60	42,078	4247	46,325	90.8	9.2
>59	11,781	1597	13,378	88.1	11.9
Area					
West-central	41,731	2075	43,806	95.3	4.7
North	5095	2996	8091	63.0	37.0
East	25,287	3115	28,402	89.0	11.0
Lexical-Morphological Affinity of the Word Form					
‘Ukrainian’	37,825	772	38,597	98.0	2.0
‘Common’	12,403	1134	13,537	91.6	8.4
‘Hybrid’	7538	394	7932	95.0	5.0
‘Russian’	14,263	5863	20,126	70.9	29.1

Table 3. **GLMM of the Influence of Area and Subcorpus (Base Level: Okannia)**

Random Effects	Variance	Std. Dev.		
Informant: Intercept	0.84	0.91		
Fixed Effects	Estimate *b*	Standard Error	*z*	*p*
Intercept	−3.52	0.14	−25.56	<0.001
Subcorpus (Reference: Family)				
Interview	0.23	0.20	1.18	0.239
Area (Reference: West-Central)				
North	3.65	0.32	11.27	<0.001
East	1.62	0.27	6.00	<0.001
Subcorpus * Area				
Interview: North	−4.11	0.53	−7.77	<0.001
Interview: East	−1.03	0.35	−2.93	0.003

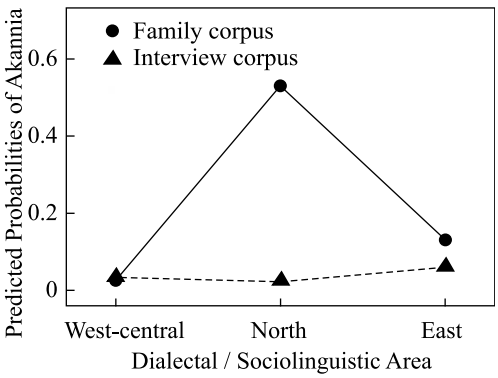

Fig. 4. Predicted Probabilities for Akannia Realizations as a Function of Area and Subcorpus

Figure 4 shows the probabilities predicted by the model for *akannia* realizations as a functionof subcorpus and area. It shows that most of the significant effects were very small, except for the effect of the northern dialectal background in the Family corpus. For most constellations, the probability of *akannia* realizations was lower than ten percent. For speakers from the eastern area in the Family corpus, the probability of *akannia* realizations exceeded 10 percent. For speakers from the northern area in the Family corpus, *okannia* and *akannia* were about equally likely.

The following analyses of the influence of further factors were carried out separately for each subcorpus.

3.3. Variation in the Family corpus

Table 4 shows the result of a GLMM for the Family corpus (number of observations: 37,901) with informant (n = 81) as a random factor.

The model reveals a general influence of gender, with *akannia* realizations being somewhat more probable for women than for men. There was no significant interaction between gender and area (χ^2 (2) = 3.77, p = 0.152). There was a significant interaction between area and age: The informants'

Table 4. **GLMM Family Corpus (Base Level: Okannia)**

Random effects	Variance	Std. Dev.		
Informant: Intercept	0.75	0.87		
Fixed Effects	Estimate *b*	Standard Error	*z*	*p*
Intercept	−3.71	0.36	−10.34	<0.001
Area (Reference: West-Central)				
North	1.80	0.82	2.19	0.029
East	2.90	0.66	4.40	<0.001
Age (in 2014)	−0.00	0.01	−0.17	0.866
Gender (Ref: Male)				
Female	0.45	0.23	1.97	0.049
Area * Age				
North: Age	0.05	0.02	2.42	0.016
East: Age	−0.03	0.01	−2.25	0.024

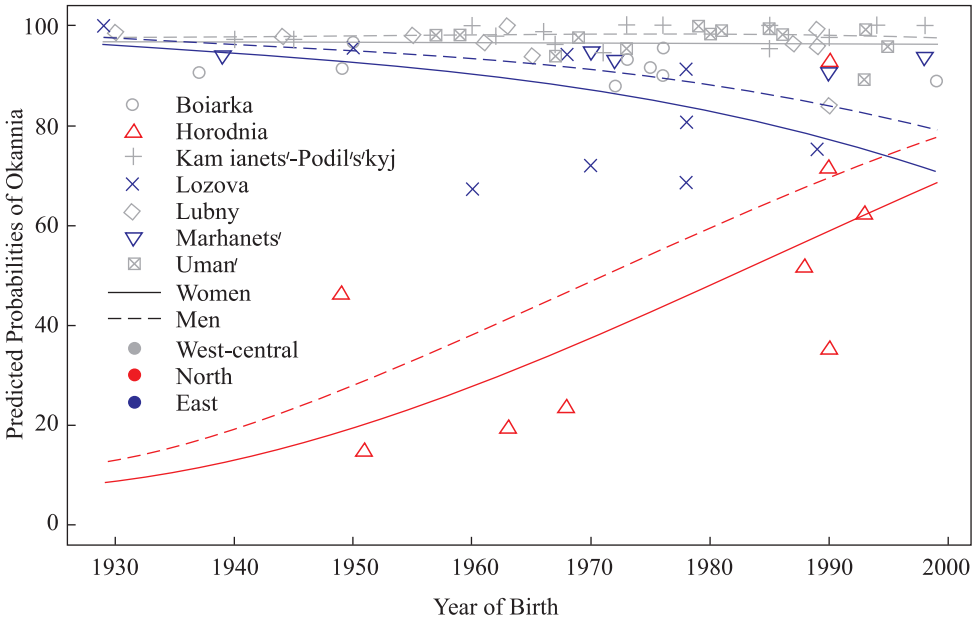

Fig. 5. Effects of Age, Gender, and Area in the Family Corpus

age had an effect on their pronunciation of the vowel variable, but this effect varied between the different dialectal/sociolinguistic areas. This is shown in Figure 5.

Informants from the west-central area almost exclusively produced realizations of *okannia*, regardless of their age or gender (grey line). This is also true for the oldest respondents from the eastern area (blue line). Middle-aged and younger respondents from this area exhibited a slight trend towards more *akannia* realizations, although the probability of *okannia* realizations remained much higher, even among the younger respondents. Women led this slight trend towards *akannia* by a small margin. Although this has not been tested, this tendency seems specific to the informants from Lozova in the Kharkiv region. The respondents from Marhanets' in the Dnipro region stuck to *okannia* pronunciation.

The tendency towards *akannia* among speakers from the eastern area was mirrored by the members of the family from Horodnia in the northern area (red line). Here, it was the older respondents who exhibited *akannia* realizations more often. Among the younger speakers from this family, the general trend was toward *okannia* realizations, although there was a great degree of variation between the speakers.

3.4. Variation in the Interview corpus

Table 5 shows the results of a GLMM for the Interview corpus (number of observations is 42,336) with informant (n = 68) as a random factor.

The only significant factor in the Interview corpus was area; however, as already shown in Figure 4, the effect of area was a small one. The factor

age did not improve the model (χ^2 (1) = 0.08, p = 0.774) — nor did gender (χ^2 (1) = 0.22, p = 0.636). Models including interactions did not converge, or the interaction was not significant.

3.5. Lexical-morphological affinity

Finally, the study examined whether the lexical-morphological affinity of the word form played a role in the variation between *okannia* and *akannia* — or, more generally, whether the phonic variation was lexically driven. This was done in separate analyses, since it can be assumed that the proportion of 'Ukrainian' and 'Russian' words in the speakers' speech also depends on the socio-demographic characteristics of the speakers. 'Hybrid' word forms

Table 5. **GLMM Interview Corpus (Base Level: Okannia)**

Random effects	Variance	Std. Dev.		
Informant: Intercept	0.68	0.82		
Fixed Effects	Estimate *b*	Standard Error	*z*	*p*
Intercept	−3.28	0.14	−22.97	<0.001
Area (Reference: West-Central)				
North	−0.45	0.42	−1.07	0.285
East	0.59	0.22	2.76	0.006

Table 6. **GLMM of the Influence of Lexical-Morphological Affinity, Family Corpus (Base Level: Okannia; Number of Observations: 34,551; Random Factor: Informant, n = 84)**

Random effects	Variance	Std. Dev.		
Informant: Intercept	0.72	0.85		
Fixed Effects	Estimate *b*	Standard Error	*z*	p
Intercept	−5.67	0.19	−29.20	<0.001
Area (Reference: West-Central)				
North	4.69	0.35	13.23	<0.001
East	2.53	0.30	8.34	<0.001
Affinity (Reference: 'Ukrainian'				
'Common'	2.08	0.16	12.64	<0.001
'Russian'	3.91	0.15	26.06	<0.001
Area * Affinity				
North: 'Common'	−1.25	0.20	−6.39	<0.001
East: 'Common'	−0.84	0.19	−4.42	<0.001
North: 'Russian'	−2.31	0.18	−13.17	<0.001
East: 'Russian'	−1.47	0.17	−8.62	<0.001

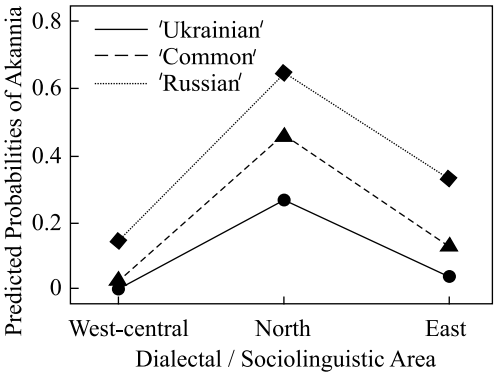

Fig. 6. Influence of Lexical-Morphological Affinity, Family Corpus

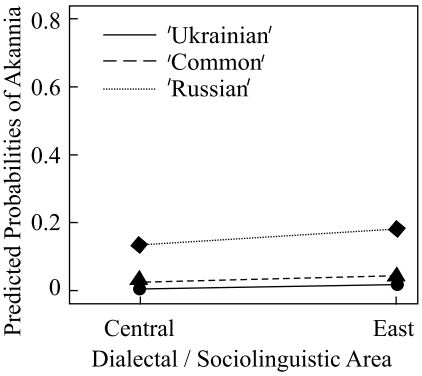

Fig. 7. Influence of Lexical-Morphological Affinity, Interview Corpus

Table 7. **GLMM of the influence of lexical-morphological affinity, Interview corpus (base level: okannia; number of observations: 37,709; random factor: informant, n = 68)**

Random effects:	Variance	Std. Dev.		
Informant: Intercept	0.48	0.69		
Fixed effects:	Estimate *b*	Standard Error	*z*	*p*
Intercept	−5.35	0.16	−33.35	<0.001
Area (Reference: West-Central or North)				
East	1.29	0.23	5.74	<0.001
Affinity (Reference: 'Ukrainian'				
'Common'	1.65	0.15	10.79	<0.001
'Russian'	3.49	0.12	28.26	<0.001
Area * Affinity				
East: 'Common'	−0.68	0.20	−3.46	<0.001
East: 'Russian'	−0.94	0.15	−6.08	<0.001

were excluded from the analyses. Tables 6 and 7 show the results of these analyses. Because no difference had been found between the western and northern areas in the Interview corpus in the previous analyses, they were not differentiated here.

In both subcorpora, the lexical-morphological affinity had an effect on the pronunciation, and this effect varied in the different subareas. The effects are shown in Figures 6 and 7.

If one ignores for the moment the northern areas in the Family corpus, word forms that — apart from their pronunciation — corresponded with the Ukrainian language almost always exhibited *okannia* realizations. It was only in the Family corpus in the northern area that there was a noteworthy probability of *akannia* realizations in 'Ukrainian' words. If one also disregards the

eastern areas in the Family corpus, the almost exclusivity of *okannia* realizations also applies to word forms that occur equally in both languages. Only in the eastern and (especially) northern area in the Family corpus, these 'common' word forms exhibited a somewhat more frequent pronunciation corresponding to Russian *akannia*. Finally, 'Russian' word forms in all constellations showed a Russian-like pronunciation more frequently than 'Ukrainian' and 'common' word forms, even if (except for the northern area in the Family corpus) *okannia* was in general dominant in 'Russian' word forms as well.

4. CONCLUSIONS

First, the results confirmed assessments by several scholars (e.g., Del Gaudio, 2010; Taranenko, 2013) that the unstressed vocalism in Ukrainian—Russian mixed speech is strongly dominated by *okannia*, i.e., the differentiation between /a/ and /ɔ/ and the phonetic realization of the latter as [ɔ]. URMS thus corresponds in this respect to the standard Ukrainian pattern and differs from Russian, where both unstressed /ɔ/ and /a/ are phonetically merged in [a]-like sounds. This is plausible in light of the emergence of URMS in a partial linguistic shift by many speakers from Ukrainian dialects towards Russian: In such situations, lexical elements of the target variety are adopted most easily, while phonic elements from the speakers' first language tend to be maintained. Nevertheless, the stability of *okannia* even among younger cohorts, who grew up in a much more Russian-influenced environment than their parents, is remarkable. There is no general trend towards a 'Russian' pronunciation, although, for many Ukrainians — at least until Ukraine's independence — Russian was the language which they had to orientate themselves towards. This is in clear contrast to the linguistic situation in Russia, where Kochetov (2006) noted a clear decline in Russian dialectal *okannia* beginning in the 1930s—1950s. Kochetov (2006, p. 115) specifically mentions the "social upheavals" of the time as an explanation for the initiation of this sound change. Such "upheavals", as is well known, were also present in the tragic history of Ukraine of that time. The presence of the standard Ukrainian language during the Soviet era, despite its discrimination in many domains, certainly contributed to the fact that a sound change from *okannia* to *akannia* either did not take place at all, or, as in the Kharkiv region, which today is linguistically Russian-dominated (Hentschel & Taranenko, 2015), only did so to a lesser extent. However, it is also remarkable for the latter region that the oldest speakers did not show any phonic influence of Russian: Even in this area that is clearly linguistically dominated by Russian today, the dialectal Ukrainian pronunciation was the starting point.

In terms of stylistic differences — or the influence of the speech situation — it is noteworthy that the Ukrainian pronunciation consistently dominated the more formal interviews, while there were some deviations from the *okannia* pattern in the informal family conversations. In formal situations, speakers clearly oriented themselves toward Ukrainian.

It is interesting to note the apparently polar opposite picture regarding the phonic variation in the north and east in the Family corpus. While there was a slight trend towards *akannia* among the speakers in the east, the opposite is true in the north. Here, a shift from *akannia* to *okannia* can be observed in general — albeit varying greatly from individual to individual. This finding must of course be interpreted with caution: There was only one family in each of these two subareas to which this statement applies, and I did not obtain any similar finding in the Interview corpus. Nevertheless, there is much to suggest that *akannia* in the northern area cannot be assessed as a Russian influence. As already mentioned, *akannia* is also present in the northern Ukrainian dialects. From this perspective, the initial polar opposite findings in the north and east turn out to be two instances of the same process: a shift from a dialectal to a standard pronunciation. In the north, this is a convergence with the standard Ukrainian pronunciation that is perhaps also assisted by the presence of *okannia* in the other Ukrainian dialects. In the east, this is a (marginal) convergence with the standard Russian pronunciation. There should be future investigations to determine whether this trend continues over time in the east or if the revaluation of the Ukrainian standard language and the Russia's war of agression against Ukraine cause a reversal.

It is unclear, however, why female respondents in both cases (i.e., in the east and north) tended to use *akannia* somewhat more often than male respondents. This suggests that the emic perspective (i.e., the evaluation of the variants from the point of view of the speakers themselves) is important in this contact situation. For the northern Ukrainian dialects, which, in some respects, correspond to patterns found in Belarusian and Russian, there is still much to be investigated from a perceptual-linguistic point of view.

Last, there are clear correlations between the phonic variation and the lexical-morphological affinity of the word form (i.e., whether the word form on the lexical-morphological level coincides with the Ukrainian or Russian language). More precisely, these correlations are co-occurrence restrictions (Kučera, 1973; Auer, 1997) that should ultimately be regarded as manifestations of an *usus* in the URMS: Not all conceivable combinations of variants are actually used by the speakers. In particular, ‘Ukrainian’ words are not pronounced in a ‘Russian’ way, whereas ‘Russian’ words allow for both ‘Ukrainian’ and ‘Russian’ pronunciation (the latter being less common). Notably, ‘common’ words behave roughly like specifically ‘Ukrainian’ words, with ‘Russian’ pronunciations rarely found. This suggests that Ukrainian is the default in an emic classification of word forms for speakers of URMS.

It does not contradict this interpretation that, in the family from the northern area in the Family corpus, ‘Ukrainian’ words quite frequently exhibit *akannia* realizations. Here, *akannia* does not mean a ‘Russian’ pronunciation of ‘Ukrainian’ words, but a ‘dialectal-Ukrainian’ pronunciation of (dialectal) ‘Ukrainian’ words. Rather, it is remarkable that ‘Russian’ (and also ‘common’) words in this family are pronounced with *akannia* more frequently than ‘Ukrainian’ words. It is possible that Ukrainian-dialectal influence and Russian

influence interact here. Another possibility is that word forms classified here as 'Russian' from an etic point of view also occur in the corresponding dialect and, accordingly from an emic point of view, are interpreted by the speakers as "own" rather than "alien" words. This should be a topic for future investigation.

To sum up, the present study represented an initial attempt to empirically investigate phonic variation in the widespread Ukrainian-Russian mixed speech ("surzhyk"). The phenomenon studied here was the variation between *okannia* and *akannia*, one of the most important phonic differences between the Ukrainian and Russian languages. The study's findings make it clear that attempts to one-dimensionally examine linguistic variation in Ukraine are bound to fail. Regional differences in terms of different dialectal starting points and sociolinguistic conditions, characteristics of the speech situation, factors such as age and gender, and, ultimately, the individual characteristics of the speakers and their families, are interconnected in a complex manner that is not easy to decipher.

REFERENCES

Auer, P. (1986). Konversationelle Standard/Dialekt-Kontinua (Code-Shifting). *Deutsche Sprache*, *14,* 97—124 (in German).

Auer, P. (1997). Co-occurrence restrictions between linguistic variables. A case for social dialectology, phonological theory and variation studies. In F. Hinskens, R. van Hout, & L. Wetzels (Eds.), *Variation, change, and phonological theory* (pp. 71—102). Amsterdam: John Benjamins.

Auer, P. (2005). Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In N. Delbecque, J. van de Auwera, & D. Geeraerts (Eds.), *Perspectives on variation. Sociolinguistic, historical, comparative* (pp. 7—42). Berlin: De Gruyter.

Avanesov, R. I. (1956). *Fonetika sovremennogo russkogo literaturnogo iazyka*. Moscow: Izdatel´stvo Moskovskogo universiteta (in Russian).

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1—48.

Bickerton, D. (1975). *Dynamics of a creole system*. London: Cambridge University Press.

Bondarko, L. V., & Verbitskaia, L. A. (Eds.). (1987). *Interferentsiia zvukovykh sistem*. Leningrad: Izdatel´stvo Leningradskogo Universiteta (in Russian).

Chertorizhskaia, T. K. (Ed.). (1988). *Ukrainsko-Russkoe dvuiazychie. Sotsiolingvisticheskii aspekt*. Kyiv: Naukova Dumka (in Russian).

Del Gaudio, S. (2010). *On the nature of suržyk. A double perspective*. Munich: Sagner (in English).

Del Gaudio, S. (2011). O variativnosti russkogo iazyka na Ukraine. *Izvestiia Rossiiskoi Akademii Nauk. Seriia Literatury i Iazyka, 70*(2), 28—36 (in Russian).

Del Gaudio, S. (2017). *An introduction to Ukrainian dialectology*. Frankfurt am Main: Peter Lang.

Del Gaudio, S. (2020). Language situation in the district of Ripky (Černihiv, Ukraine). *Russian Linguistics, 44*, 177—201.

Hentschel, G. (2014). "Trasjanka" und "Suržyk" — zum Mischen von Sprachen in Weißrussland und der Ukraine. Einführung in die Thematik und Ausblick auf den Band. In G. Hentschel, O. Taranenko, & S. Zaprudski (Eds.), *Trasjanka und Suržyk — gemischte weißrussisch-russische und ukrainisch-russische Rede. Sprachlicher Inzest in Weißrussland und der Ukraine?* (pp. 1—26). Frankfurt am Main: Peter Lang (in German).

Hentschel, G. (2018). Die weißrussische 'Trasjanka' und der ukrainische 'Suržyk'. Grundlegende quantitativ-qualitative soziolinguistische Ähnlichkeiten und Unterschiede. In S. Kempgen, M. Wingender, & L. Udolph (Eds.), *Deutsche Beiträge zum 16. Internationalen Slavistenkongress Belgrad 2018* (pp. 127—138). Wiesbaden: Harrassowitz (in German).

Hentschel, G., & Reuther, T. (2020). Ukrainisch-russisches und russisch-ukrainisches Code-Mixing. Untersuchungen in drei Regionen im Süden der Ukraine. *Colloquium: New Philologies, 5*(2), 105—132 (in German).

Hentschel, G., & Taranenko, O. (2015). Die Sprachenlandschaft der zentralen Ukraine: Ukrainisch, Russisch, 'Suržyk'. Verwendung — Kompetenz — nationale Positionierung. *Die Welt der Slaven, 60*(2), 248—275 (in German).

Hentschel, G., & Zeller, J. P. (2012). Gemischte Rede, gemischter Diskurs, Sprechertypen. Weißrussisch, Russisch und gemischte Rede in der Kommunikation weißrussischer Familien. *Wiener Slawistischer Almanach, 70*, 127—155 (in German).

Hentschel, G., & Zeller, J. P. (2017). Aspekte der Sprachverwendung in zentralen Regionen der Ukraine. *Wiener Slawistischer Almanach, 79*, 37—60 (in German).

Hryshchenko, A. P., Mats′ko, L. I., Pliushch, M. Ia., Tots′ka, N. I., & Uzdyhan, I. M. (2002). *Suchasna ukraïns'ka literaturna mova*. Kyiv: Vyshcha shkola (in Ukrainian).

Kasatkin, L. L. (2006). *Sovremennyi russkii iazyk. Fonetika*. Moscow: Akademiia (in Russian).

Kerswill, P. (2007). Koineization and accommodation. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 669—702). Malden, MA: Blackwell Publishers.

Kochetov, A. (2006). The role of social factors in the dynamics of sound change. A case study of a Russian dialect. *Language Variation and Change, 18*(1), 99—119.

Kučera, H. (1973). Language variability, rule interdependency, and the grammar of Czech. *Linguistic Inquiry, 4*, 499—521.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Le Page, R. B., & Tabouret-Keller, A. (1985). *Acts of identity. Creole-Based approaches to language and ethnicity*. Cambridge: Cambridge University Press.

Panov, M. V. (1979). *Sovremennyi russkii iazyk. Fonetika*. Moscow: Vysshaia shkola (in Russian).

Pompino-Marschall, B., Steriopolo, E., & Żygis, M. (2017). Ukrainian. *Journal of the International Phonetic Association, 47*(3), 349—357.

R Core Team (2021). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*. https://www.R-project.org/

Rusanovskii, V. M., Zhovtobriukh, M. A., Gorodenskaia, E. G., & Grishchenko, A. A. (1986). *Ukrainskaia grammatika*. Kyiv: Naukova Dumka (in Russian).

Schilling-Estes, N. (2002). Investigating stylistic variation. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 375—401). Malden: Blackwell.

Šumarova, N. (2014). Der Suržyk im System nah verwandter Zweisprachigkeit. Soziolinguistischer und linguistischer Aspekt. In G. Hentschel, O. Taranenko, & S. Zaprudski (Eds.), *Trasjanka und Suržyk — gemischte weißrussisch-russische und ukrainisch-russische Rede. Sprachlicher Inzest in Weißrussland und der Ukraine?* (pp. 307—326). Frankfurt am Main: Peter Lang (in German).

Taranenko, O. (2013). Variantnist′ vs. stabil′nist′ u strukturi ukraïns′ko-rosiis′koho "surzhyku." Sukupnist′ idiolektiv vs. sotsiolekt. In G. Hentschel (Ed.), *Variation und Stabilität in Kontaktvarietäten. Beobachtungen zu gemischten Formen der Rede in Weißrussland, der Ukraine und Schlesien* (pp. 27—61). Oldenburg: BIS (in Ukrainian).

Timberlake, A. (2004). *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Trub, V. M. (2000). Iavyshche "surzhyku" iak forma prostorichchia v sytuatsiï dvomovnosti. *Movoznavstvo, 1*, 46—58 (in Ukrainian).

Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.

Voss, Ch. (2008a). Slawische Kreolsprachen — Mythos und Realität. In S. Kempgen, K. Gutschmidt, U. Jekutsch, & L. Udolph (Eds.), *Deutsche Beiträge zum 14. Internationalen Slavistenkongress Ohrid 2008* (pp. 357—369). Munich: Sagner (in German).

Voss, Ch. (2008b). Zur Konzeptualisierung sprachlicher Hybridität in der Slawia. "Wasserpolnisch" in Schlesien und "Suržyk" in der Ukraine. In A. Nagórko, S. Heyl, & E. Graf (Eds.), *Sprache und Gesellschaft. Festschrift für Wolfgang Gladrow* (pp. 543—557). Frankfurt am Main: Peter Lang (in German).

Watt, D., & Milroy, L. (1999). Patterns of variation and change in three Newcastle vowels. Is this dialect levelling? In P. Foulkes & G. J. Docherty (Eds.), *Urban voices. Accent studies in the British Isles* (pp. 25—46). London: Arnold.

Wickham, H. (2019). *stringr: Simple, consistent wrappers for common string operations.* R package version 1.4.0. https://CRAN.R-project.org/package=stringr

Yanushevskaya, I., & Bunčić, D. (2015). Russian. *Journal of the International Phonetic Association, 45*(2), 221—228.

Zeller, J. P. (2018). Zum ukrainisch-russischen Sprachkontakt: Phonische Variation im ukrainischen 'Suržyk' im Vergleich mit der weißrussischen 'Trasjanka'. In S. Kempgen, M. Wingender, & L. Udolph (Eds.), *Deutsche Beiträge zum 16. Internationalen Slavistenkongress Belgrad 2018* (pp. 365—374). Wiesbaden: Harrassowitz (in German).

Zeller, J. P., Taranenko, O., & Hentschel, G. (2019). Language and religion in central Ukraine. *International Journal of the Sociology of Language, 2019*(260), 105—130.

Zhylko, F. T. (1966). *Narysy z dialektolohiï ukraïns'koï movy.* Kyiv: Radians'ka shkola (in Ukrainian).

Received 28.03.2022
Accepted 15.05.2022

Ян Патрік Целлер, Dr., професор кафедри слов'янської лінгвістики,
Ґрайфсвальдський університет, Німеччина
Ernst-Lohmeyer-Platz 3, 17487 Greifswald
Tel.: +49 (0)3834 420 3222
E-mail: jan.zeller@uni-greifswald.de
https://orcid.org/0000-0002-6942-7112

ОКАННЯ І *АКАННЯ* В УКРАЇНСЬКО-РОСІЙСЬКОМУ ЗМІШАНОМУ МОВЛЕННІ («СУРЖИК»)

Попри те, що українсько-російське змішане мовлення (УРЗМ) є поширеним явищем у центральних регіонах України, емпіричних досліджень щодо варіації звукових характеристик УРЗМ та зв'язку з соціальними характеристиками його мовців бракує. Спираючись на корпус розмовного мовлення (близько 340000 словоформ), створений зі спонтанних неформальних розмов та відкритих інтерв'ю, автор аналізує варіювання між двома яскравими звуковими властивостями української та російської мов, що відбиті в УРЗМ, а саме ненаголошеним /ɔ/ в українському *оканні* та російському *аканні*. Результати дослідження підтверджують, що в ненаголошеному вокалізмі УРЗМ істотно домінує *окання*, яке відповідає стандартному українському зразку, відмінному від російського, де ненаголошені /ɔ/ та /a/ фонетично злилися у звукові [а]-типу. Узагальнене лінійно-змішане моделювання засвідчує, що варіація між *оканням* та *аканням* пов'язана з такими чинниками, як діалектні та соціолінгвістичні відмінності в українському мовному ландшафті, соціально-демографічні характеристики мовців, як-от вік і стать, а також характеристики мовленнєвої ситуації. Значення має кореляція між фонетичною варіацією та лексико-морфологічною спорідненістю словоформи, тобто чи на лексико-морфологічному рівні словоформа відповідає українській чи російській системі.

Ключові слова: українсько-російські мовні контакти, діалектні контакти, мовна варіативність, окання, акання