

On vocatives and discourse markers in Ukrainian *Suržyk*: Ukrainian or Russian?

Abstract

The guiding question of this study is whether syntactically isolated expressions with a discourse-organizing function exhibit a systematic preference for Russian variants over their functionally equivalent Ukrainian counterparts in the mixed code of Ukrainian *Suržyk*. The underlying assumption is that syntactically isolated expressions are more readily transferred from a socially dominant language (in Ukraine, Russian for more than a century) into the speech of speakers of a dominated language (Ukrainian, at least until 1990). With “syntactically isolated expressions” we refer to vocatives and discourse markers.

A corpus-based analysis was conducted to determine the frequencies of three types of vocative expression: the (in Ukrainian) inherited vocative case, an endless “neo-vocative,” and the nominative. For discourse markers, 16 Ukrainian-Russian “hyperlexematic” pairs encompassing 70 functional equivalents from both donor languages, Ukrainian and Russian, with 17,000 tokens were analyzed. While the distribution of variants expressing vocative functions appears, at first glance, to support the hypothesis formulated above, the results for discourse markers do not. Rather, discourse markers display a redistribution of Ukrainian and Russian “translational equivalents” similar to that observed for “lexically full” parts of speech such as nouns, verbs, and adjectives. In some cases, the Ukrainian variants, in other cases the Russian variants are firmly established in *Suržyk*. A third group of discourse markers is characterized by variation between Ukrainian and Russian expressions. This variation is, if at all, largely determined by the sociobiographical factor of the speakers’ exposure to Russian in everyday interaction and shows only marginally regional differentiation, in contrast to the variation amongst Ukrainian and Russian nouns, adjectives, and verbs, where regional factors play a clearly dominant role.

1 Introduction

The guiding question for this study is whether syntactically isolated expressions with an organizing function in discourse in the mixed code of Ukrainian *Suržyk* exhibit a specific tendency to be realized with Russian variants rather than with their functionally equivalent Ukrainian counterparts. This question—or, more precisely, the underlying hypothesis—is based on two “pillars”.

First, the mixed code of Ukrainian *Suržyk*, with elements and constructions from Ukrainian and Russian, developed in the context of the social and political dominance of Russian in Ukraine, which lasted for more than a century (cf. Taranenko, 2007). This “traditional” *Suržyk* is clearly Ukrainian-based, in the sense that speakers of Ukrainian, mainly of dialectal varieties, tried to *accommodate* (in the sense of Trudgill, 1986) to Russian. In any case, the impact of Russian has been considerable, at least until the end of the Soviet Union. Even today this impact is highly visible, not only in the lexicon (cf. Hentschel, 2024), but also in rather closed linguistic subsystems, such as inflectional morphology (cf. Hentschel, forthcoming a, forthcoming b) or in sound patterns (cf. Zeller, forthcoming). Whether, in the roughly 35 years of independent Ukraine—during which the social position of Ukrainian has been consistently strengthened politically—this Ukrainian-based *Suržyk* has changed in a way that it has gradually again become more Ukrainian—at least with the younger generation—cannot be clearly determined. The papers in this special issue are based on linguistic material from a *Suržyk* of the last 10 years (cf. Hentschel et al., forthcoming). There is, however, no basis to compare today’s *Suržyk* with that around the year 1990 due to the complete lack of solid empirical investigations on *Suržyk* at that time. Nevertheless, neither the studies in this volume nor other recent ones such as Hentschel (2024), Hentschel and Palinska (2022) or Palinska and Hentschel (2022) indicate a broad influence of age such that the *Suržyk* of the older generation is more “Russian-like” than that of a generation linguistically socialized in independent Ukraine (cf. Hentschel, forthcoming a). In other words, from the “apparent time point of view”, asking for possible differences between speakers from different generations, *Suržyk* appears to be relatively stable.

A further question, namely whether a new, Russian-based variety of *Suržyk* could have developed, is discussed in this special issue by Hentschel (forthcoming a). He argues that there is no empirical evidence for a Russian-based *Suržyk* and furthermore that it could hardly have developed within 35 years into a relatively stable “low-variety,” a designation appropriate for Ukrainian-based *Suržyk*, which itself developed over a time span of more than 100 years. The linguistic material on which the following analysis is based must be classified as Ukrainian-based *Suržyk*.

Second, the question was motivated by the assumption that linguistic expressions in syntactic isolation, such as vocatives and discourse markers, are transferred much more easily from a dominant superstrate into a dominated substrate than syntactically integrated elements, in a type of linguistic behavior that Trudgill (1986) calls *imitation*—even when speakers of the substrate have only a rather weak command of the superstrate.

Code-mixing in Suržyk can generally be characterized as the type that Muysken (2000) calls *congruent lexicalization*, which is typical of contact situations between structurally and genetically related codes (cf., for example, Hentschel, forthcoming b on inflectional morphology). This does not rule out the possibility that, while speaking, Suržyk speakers switch to (Standard¹) Ukrainian or (Standard) Russian, or gradually slide toward the standard varieties, if the situation rather demands “pure” speech. The extent of the approximation toward the standards depends on the speaker’s command of them.

2 Introductory theoretical remarks

Vocatives and discourse markers have—among other things—in common that they occur in syntactic isolation, in some kind of dislocation. This isolation is the basis for their discourse-organizing function. The term *discourse marker* will be used here in a broad sense. In the literature, especially in older publications, the expressions analyzed below are often classified as interjections or particles. Whether they are restricted to positions outside sentence structure is not only an empirical question but a theoretical and terminological one as well. For example, Isačenko (1975, pp. 629–632), half a century ago, stated with regard to interjections that it would be false to assume that they occur exclusively in a syntactically isolated position. He speaks, for example, of an *interjection* in the case of Russ. *dejstvitel’no* ‘really’ (Ukr. *dijsno*). However, in a sentence like *Ona dejstvitel’no ne ljubit krasnoe vino* ‘She really does not like red wine’ the traditional part of speech assumed for *dejstvitel’no* would be the adverb. However, if *dejstvitel’no* were an isolated expression of confirmation of an immediately preceding utterance or an expression asking for confirmation of such an utterance, it would rather be classified as a particle (Gorbačevič, 2004–, s.v.). Similarly, forms of the second person present tense of certain verbs are widely described as discourse markers only when they occur in an isolated, dislocated position, be it to the right, to the left, or somewhere within a “complete” sentence structure—cf. for example sentences (1) and (2) from the *Oldenburg corpus of Ukrainian Suržyk* (hereafter OCUS) this study is based on (cf. Hentschel et al., forthcoming and, in more detail, the documentation of the corpus by Hentschel et al., 2026, hereafter OCUS-Doc²):

- (1) *A tak my zaras piŭ torta zjimo, ponimaješ, čaju³ napjemos’.*
‘So now we’ll eat half the cake, you understand, and we’ll have some tea.’⁴
- (2) *My, rozumiješ, nikoly ne spišymo, i ty ne spišy!*
‘We, you understand, never hurry, and you don’t hurry either!’⁵

In sentences where, on the other hand, the same verb forms govern a nominal direct object or sentential object, they are, of course, described as “normal” syntactic verbal predicates and not as verbs. It suffices here to hint at translations with English sentences like *You don’t understand my situation* or *You probably understand that he had to leave earlier*.

¹ When we refer in brackets to the spoken “Standards” of Ukrainian or Russian, we mean the colloquial usage of the standardized, codified varieties. This speech may contain regional features, but no current instances of Russian influence.

² The documentation is available at <https://uol.de/slavistik/forschung/sprachwissenschaft/bereich-mischvarietaeten/okus-dok>.

³ Note that citations of material from the corpus are given here in a specific transcription that resembles the usual transliteration (cf. OCUS-Doc). The *j* after *č* (in Cyrillic alphabet *ч*) in the OCUS transcriptions signals the palatalized, fronted or “soft” realization of the affricate, which is the norm in Russian. The Ukrainian norm is a “hard” articulation, moved somewhat back to the velum. However, as Zeller (forthcoming) reports, the former is the clear usage norm in Suržyk, too.

⁴ Please note that our translations are, as far as possible, rather literal, and thus stylistically often not the best.

⁵ In Ukrainian and Russian, verb forms in this function are realized without the corresponding second person pronoun as the “subject”. In Germanic languages such as German or English, uses of corresponding verbs are realized as sequences of deictic pronoun and the verb form, cf. Engl. *you understand* or Germ. *du verstehst* or inverted as a question *verstehst du?*

Turning to vocatives, it is well known that these specific morphologically marked forms that are used to directly address a person (or personalized referent) can often be replaced in discourse by nominative forms. These nominatives in vocative function are usually not called vocatives in a syntactically isolated position or vocative (addressing) function. Nominative forms in vocative function are widespread in Slavic languages that contain specific vocative forms. On the other hand, modern Slavic standard languages that contain such specific morphological vocative forms in their inflectional inventories generally do not exhibit the phenomenon of vocative forms in the function of the (surface) subject in sentences “substituting” the nominative. In historical or non-standard Slavic variants this, however, is a well-known phenomenon (cf. Stifter, 2013, pp. 61–67). The terminological tradition, however, is inconsistent. Nominative forms in vocative function are consistently called nominatives, vocative forms as subjects “remain” vocatives. Stifter (2013, pp. 61–67) speaks of “vocatives for nominatives”. We will use *vocatives* as a general expression for elements in vocative function and differentiate between several subtypes if not explicitly stated otherwise.

Syntactic isolation will thus be a decisive criterion for an expression to be considered in the analyses. It makes the expressions discussed here interesting for contact linguistics. Trudgill (1986) differentiates two transfer processes from one dialect or variant of a language to another:⁶ *Accommodation* and *imitation*. Roughly speaking, accommodation is more of a long-term unconscious process. The use of a given expression E from donor code A in receiving code B would be based on a rather solid establishment of E in B. Imitation, however, is a rather spontaneous, intentional use of an expression E from code A by a speaker of dialect B imitating A, the latter generally being the variant of higher (overt or covert) social prestige in the current surroundings.

One should rather think of the differentiation between *accommodation* and *imitation* in gradual terms, where imitation may appear at the very beginning of a contact situation when speakers of B are expected, if not forced to adapt to A. Imitation may thus give birth to accommodation, without being abandoned as accommodation progresses. As, for example, Chambers (2002) has pointed out, speakers have a certain competence in judging what share of elements from variants or registers of a language is appropriate in specific conversational settings. Expressions that are not bound into a syntactic construction are much easier for a speaker of the code B, with its lower social prestige, to employ in an imitational way in conversational settings where the use of more expressions of code A, with its higher prestige, seems appropriate. In Poplack’s (1980, p. 614) terms one might speak here of an *emblematic usage* of the corresponding discourse-organizing elements.

Turning to Suržyk, imitating Russian was particularly relevant as long as Russian was the language of power in Ukraine. This constellation was conducive to an imitation-based usage of Russian words. In independent Ukraine, the situation has changed to a considerable degree. Now, conversational settings in which Ukrainian is expected, or at least “linguistic closeness” to Ukrainian is desirable, are much more frequent and thus the “imitational attractiveness” of Ukrainian discourse markers and similar elements has increased.

Imitation that turns into accommodation may thus support the transfer or borrowing of corresponding expressions from A to B. This leads the discussion to borrowing hierarchies, at least for discourse markers and interjections. Vocatives, here in the sense of specific inflectional forms of nouns and proper names, have so far, to the best of our knowledge, only marginally been examined in relation to borrowing hierarchies.

Matras (2007, 2009, pp. 153–161) discusses several variants of borrowing (borrowability) hierarchies that have been proposed by linguists.⁷ He himself places discourse markers (presumably in a narrower sense) as well as interjections relatively high on a cross-linguistic scale, which is not (directly?) based on the quantity of tokens or types, but “rather on the number of languages in the [his] sample showing borrowings of the relevant categories” (cf. Matras, 2009, p. 157):

nouns, conjunctions > verbs > *discourse markers* > adjectives > *interjections* > adverbs > other *particles*,
adpositions > numerals > pronouns > derivational affixes > inflectional affixes

Haugen (1950, p. 224), analyzing Norwegian and Swedish immigrant speech in the United States, on the other hand, places interjections at the end of his hierarchy, lower than the main parts of speech comprising nouns,

⁶ Transfers from Russian to a Ukrainian-based Suržyk or to Ukrainian itself structurally function like transfers between dialects of one language, which is due to the genetic and structural closeness between the two languages (cf. Thomason & Kaufman, 1988).

⁷ Hentschel (2018) shows that in the development of Suržyk, borrowing hierarchies are principally relevant, empirically stating a hierarchical ordering of transfers *nouns* > *verbs* > *pronouns* from Russian into a Ukrainian-based Suržyk and its Belarusian equivalent, the Belarusian-based Trasjanka.

verbs, and adjectives and together with adverbs and prepositions. At Haugens's time, the term *discourse marker* was not yet established in linguistic terminology.

Andersen (2014) points out that discourse-organizing elements contradict the traditional view on function words, since they are quite easily transferred between languages, more than other function words. He postulates the following partial borrowability hierarchy, even including vocatives:

interjections/response particles > discourse markers > vocatives > hesitation markers > other function words

Andersen's list, in comparison to the others cited, suggests that—apart from the question of hierarchical ordering—different students of pragmatic and discourse-organizing expressions use different definitions or delimitations of relevant parts of speech: discourse markers (in a narrower sense), interjections, particles, etc.

More recent studies, such as those by Peterson (2017), Mišić Ilić (2017), and Unuabonah and Mabena (2024), rather underline that the borrowability of discourse markers depends on the subtype in a differentiated categorization of discourse markers. Furthermore, factors like syntactic non-integration, a higher degree of desemantization, a high social indexicality (urbanity, social position, youth style, etc., which tend to not be considered in typological studies) support borrowability. Factors like these suggest that it may be highly problematic to establish a universal hierarchy of the parts of speech and the expressions this study is going to concentrate on, because the social conditions depend very much on the concrete “landscape of linguistic contact” with its historical, political, and social implications for the contacting languages or codes in question.

In this paper, which will be based on a quantitative analysis of data in an electronic corpus of (the almost exclusively spoken) Suržyk, OCUS, we cannot and will not go into a theoretical discussion on delimiting the parts of speech of relevance or on possible subtypes within them. There is a rich literature on these questions (cf. for examples Norrick, 2007, 2009; Theodoropoulou, 2009; Proske, 2014; Dingemanse, 2024).⁸ The expressions to be analyzed below are firstly all widely described as either discourse markers, interjections, particles, or as specific vocative forms, when, secondly, they occur in syntactic isolation.

3 Further methodological clarifications

There are further conditions for selecting the expressions to be analyzed (or not). We do not consider expressions that Isačenko (1975) calls “Lautreflexe” (sound reflexes) such as Russ. *a!*, *o!*, *ax!* *ox!* *aj!* *oj!* These are sometimes called primary interjections. Close to them are response tokens like Russ. *aga*, *uvy* expressing, for example, doubt, regret and the like. Interjective, vocative uses of deictic second person pronouns and swear words as in *Ty, svoloč'*, *zatknis'*, literally: ‘You, scumbag, shut up!’ are neglected, too. In many cases, corresponding forms in Ukrainian and Russian would be homonymous. Homonymous expressions in Ukrainian and Russian are excluded on principle, as they do not provide any insights into the respective impact of Ukrainian and Russian on Suržyk.

The study is corpus-based. Like all (but one) other papers in this special issue of *Russian Linguistics* this study exploits OCUS. As has been said above, this material reflects the mixed speech of the “traditional” Ukrainian-based Suržyk.⁹ This corpus contains about 730,000 word forms. Other papers in the special volume that analyze the corpus material concentrate on hybrid utterances, which *per definitionem* include Ukrainian and Russian lexical or grammatical words and (smaller) elements or constructions, alongside common ones. These hybrid utterances constitute about 530,000 word forms. Ukrainian utterances (with Ukrainian or common characteristics) encompass approximately 150,000 word forms, Russian ones (with Russian or common characteristics) 33,000, and common ones, which represent possible utterances in either language (notwithstanding “surface” phonetic traits) comprise 8,000. These figures, too, reflect Hentschel's (forthcoming a) point that the material represents Ukrainian-based Suržyk: Hybrid utterances alternate with Ukrainian ones in discourse with a relation of roughly 3.5 to 1, whereas the relation between hybrid and Russian utterances is 16 to 1 and hybrid to common utterances is 66 to 1.

⁸ For an overview of approaches and terminology in the German and Slavic linguistic traditions, see Blankenhorn (2003, pp. 77–152).

⁹ The question of an alleged Russian-based Suržyk or Neo-Suržyk is discussed in Hentschel (forthcoming a).

Restricting the analysis to hybrid utterances is reasonable for morphological and lexical studies that focus on elements that are integrated into sentence structures. The focus on hybrid sentences represents an analytically safer way to shed light on the question of whether we are dealing with code-mixing or with a fused lect in the sense of Auer (1999) by concentrating on “inter-sentential” phenomena.¹⁰ It would, however, have been counterproductive if we had followed this restriction in our analyses here: Many of the expressions of interest here occur in one-word utterances. In the overwhelming majority of instances, this one word (form) is either common to both Ukrainian and Russian and thus of no interest for measuring the Ukrainian and Russian impact on Suržyk, or it can be unambiguously classified as Ukrainian or Russian. Therefore, we analyze all the material. It should be noted that hybrid utterances encompass more than 70% of the corpus material, so it is obvious that the speakers rarely use a “pure” Ukrainian or Russian code in the material recorded.

4 Analyses

4.1 Vocatives

4.1.1 Forms in Ukrainian and Russian

As is well known, there are diverging opinions on whether specific inflectional forms of personal names or nouns that are functionally vocatives should be considered as forms of a corresponding (morphological or surface) case. With regard to the reflexes of the Old Slavic morphological vocative, this question would however be relevant only for Ukrainian. It is widely agreed that the very few reflexes in (Standard) Russian such as *Bože* ‘God’, *Gospodi* ‘Lord’, and *otče* ‘Father’ should be qualified as interjections and not as case forms. They may be used in the primary vocative function as an invocation of the Omnipotent (e.g., in a prayer) or, in the case of *otče*, to address a priest. In very many cases they are used as an expression of astonishment, fright, horror, or similar emotions (cf. Zwicky, 1974, p. 787 for the scope of such functions).

The corresponding forms in Ukrainian (*Bože*, *Hospody*, *otče*) and several other specific vocative forms (e.g., *mamo*, primarily a call for or address to one’s mother) can be used as interjections, too. However, at least at first glance, they correspond to a still productive pattern of building vocative forms (see below). All in all, however, this question whether the vocative (in Ukrainian) is a case, is of no importance for our analysis and potentially not even for the phenomenon in general (Sonnenhauser & Hanna, 2013b, pp. 4–7). Following the established tradition for Ukrainian, we occasionally speak of vocative *case* forms below, which is motivated by inflectional paradigms.

The vocative function of inflected nouns in Suržyk (as documented in OCUS) is expressed by the following forms:

- I. *Nominative forms in vocative function* (NOMVOC). These are the normative means of expressing vocative function in Standard Russian. The historical morphological vocative, different from the nominative case, has not been preserved in Russian (see above and point II). For Ukrainian, due to the lack of research on spoken, everyday Ukrainian, it is hard to quantify the extent to which normative morphological vocative forms have been replaced by a nominative in the spoken standard variety.¹¹ However, earlier studies leave no doubt that such replacements are widespread in Ukrainian speech (e.g., Shvedova & Lukashevskyj, 2025 for the last decades, see below).¹² The nominative in vocative

¹⁰ In casual speech, not all utterances are structurally “complete” sentence, especially one-word utterances. Most are, especially those containing more than one or two word forms.

¹¹ Unfortunately, we cannot take advantage of the Ukrainian GRAK corpus as a baseline for comparison. First, the corpus consists predominantly of written texts, in which the use of the traditional vocative is likely to be more consistent, especially in edited texts. One of the anonymous reviewers of this paper cites Oleksa Nehrebetskyj, a well-known translator and film-dubbing director, stating that he hates editors who force him to use specific vocative forms even with foreign names and sometimes with ridiculous results, e.g., transferring an Engl. *Sir Pole* into a Ukr. *Sere Pole*, Ukr. *sere* = ‘shit: 3. person singular present’, *pole* = ‘field’. Second, GRAK does not (yet) include systematic grammatical homonymy resolution in the sense that nominative forms in vocative function are specifically annotated. Cf. <https://theukrainians.org/oleksa-nehrebetskyi/>.

¹² In Polish, where there is a specific vocative form as in Ukrainian, the nominative instead of a vocative is often a more informal, if not coarser form of address. For example, after some unsuccessful attempts to address a person with a vocative form, a rather harsh nominative form may follow: *Andrzeju*, ..., *Andrzeju*, ... [nominative, louder:] *Andrzej!* *Co jest, śpisz?* Which in English would be: *Andrew*, ... *Andrew* ..., *Andrew!* *What’s going on? Are you sleeping?*

function—usually marked in some grammars as “less common”—is described in normative grammars as a trait of Russian influence (for a detailed history of normative changes in Ukrainian in the 1930s and later, see Kobčenko, 2021). The process of partially replacing the vocative with the nominative is also noted by Ju. Ševel’ov (Šerex, 1951, p. 233), who, however, regards it as part of a general trend toward a reduction in the number of case forms in Ukrainian, i.e., not necessarily as a result of Russian influence. Furthermore, in Ukrainian dialects, nominative forms in vocative function seem to be somewhat more widespread than is assumed for the standard language (Bevzenko et al., 1978, p. 82). However, only spoken (approximate Standard) Ukrainian of “everyday people”¹³ would be relevant for a comparison with dialects. We know little about the frequency of nominatives in vocative function in the latter.¹⁴

- II. *Ukrainian-specific morphologically overt vocative forms continuing the historical (traditional) morphological Old Slavic vocative: (TRADVOC).* Vocative forms in Ukrainian are structurally possible for masculine and feminine nouns. For neuter nouns and all plural forms, there are no specific forms, except for a few neuter exceptions, e.g., *imene* ← *im* “ja” ‘(first) name’, *plemene* ← *plem* “ja” ‘tribe’ (Mazničenko et al., 2019, p. 131), and plural vocative forms such as *panove* ← *pany* ‘gentlemen’, *bratove* ← *braty* ‘brothers’, and *synove* ← *syny* ‘sons’ in poetic and solemn speech (Mazničenko et al., 2019, p. 128).¹⁵ In all other cases, nominative forms are used in the vocative function for the neuter singular and for all plural forms. Adjectives in Ukrainian do not possess specific vocative forms and are therefore used exclusively in the nominative in functionally vocative constructions like noun phrases containing nouns in the specific vocative forms.

Traditional vocative forms are also typical for most Ukrainian dialects, which means that they are in principle present in oral variants of Ukrainian, too. The extent of its dialectal usage today, however, is not very clear (see below).

If we think of specific vocative forms as markers of case, then this vocative has most probably always been a *facultative* case in Slavic languages (when spoken), in the sense of Zaliznjak (1973, pp. 80–84), whereby the degree of *facultativity* certainly varies between languages and, within a given language, between different varieties or registers and across different times. The degree of facultativity is rather dependent on sociolinguistic norms, and is not a systemic question.

Regarding the traditional vocative forms themselves, in regard to the analysis to follow, it suffices to note that they are marked by the endings /-o/, /-e/, or /-u/, depending on the inflectional class and/or the final consonant of the stem; some examples (nom. → voc.): *Oksana* → *Oksano*, *Katja* → *Katju*, *Marija* → *Marije*, *Ivan* → *Ivane*, *Andrij* → *Andriju* (all personal names), *suddja* → *sudde* ‘judge’, *druh* → *družē* ‘friend’, *xlopčyk* → *xlopčyku* ‘boy’. If the stem final consonant is a velar, then, before /-e/, this velar turns into a postalveolar affricate or fricative (*družē*). This reflects an Old Slavic sound change in the second half of the first millennium after Christ. Not all vocative forms are morphologically specific, though always different from nominative forms. Especially the forms on /-u/ coincide either with accusative (mostly for feminine gender) or dative (masculine) forms.

- III. *Forms with a “loss” of final vowel, i.e., the ending of the nominative, the so-called “neo-vocative” (NEOVOC).* It should be noted that such forms, bare stems, occur already in Russian literature of the 19th century in characters’ direct speech. Initially, they were associated primarily with speakers of lower social status (e.g., peasants and sailors), but from the 1920s onward they spread to other social groups as well. These forms are widespread in contemporary, colloquial Russian (cf. Daniel, 2009, Janda 2019). There are some formal constraints in this group: the overwhelming majority (see below) of these forms are derived from historical so-called *a*-declension nouns, with a nominative singular on /-a/, primarily kinship terms (*mam* ← *mama*, *pap* ← *papa* ‘mummy/daddy’) and personal names (*Saš* ← *Saša*, the latter an affectionate form for *Aleksandr*). With regard to non-monosyllabic stems, according to Daniel (2009), based on research on the Russian National Corpus (NKRJa)¹⁶, neo-vocatives are preferably formed from words with penultimate stress (*ba’bul’* ← *ba’bulja*, rarely *’babuška* ← *’babuška* ‘grandma’). Furthermore, a single final consonant supports the formation: thus, *Saš* ← *Saša* is much more common than *Sašk* ← *Saška*; *Sašk* is perceived as rather vulgar (“prostorečie”). Lexical-pragmatic restrictions may apply: it is problematic to form neo-vocatives for words that are rarely used in the address function, e.g., *’těšč* ← *těšča* ‘mother-in-law’, *’nevest* ← *nevesta* ‘bride’, *’mačex* ← *mačexa* ‘stepmother’. Typically, neo-vocatives are not formed for principally connotated words like *suka*

¹³ Some observations can be found in Taranenko (2024, pp. 97–104) on the use of the nominative in the function of the vocative, primarily regarding tendencies at the end of the 20th and beginning of the 21st century, related to processes of so-called democratization and liberalization in the spheres of language use and language regulation.

¹⁴ Bevzenko (1980, p. 95) states that in most Polesian dialects, including the Left-Bank Polesian and Right-Bank Polesian varieties, vocative forms coincide with the nominative case, i.e., there are no specific vocative forms.

¹⁵ Matvijas (1974, p. 122), describes these forms with -ove as archaic residues of the nominative plural of ancient *ŭ*-stems.

¹⁶ Cf. Savčuk et al. (2024).

‘bitch’, *duška* ‘sweetheart’, *sestrica* ‘sister (diminutive)’ (Daniel’, 2009, pp. 225–229). On the other hand, even plural forms can occur when other formal conditions are matched: *rebjat* ← *rebjata* ‘young folks’, *devčat* ← *devčata* ‘girls’.¹⁷

Ukrainian linguistics either remains silent on the matter of NEOVOC in Ukrainian or unequivocally classifies such forms as elements of Suržyk or, more generally, as the result of Russian influence. In particular, Taranenko (2024, pp. 97–104) perceives neo-vocative forms as Russian forms, though they are widely attested in the *General’nyj rehional’no anotovanyj korpus ukraińs’koï movy* (GRAK) and in Ukrainian dialects (cf. Bevzenko 1980, p. 95).

- IV. *Forms with the final syllable truncated* (SYLTRVOC): e.g., *xlo* ← *xlopec* ‘boy’, *dja* ← *djadja* ‘uncle’, which Taranenko (2024, pp. 97–104) principally describes as typical for Ukrainian (without interference from Russian). They are indeed widespread in Ukrainian dialects (again Bevzenko 1980, p. 95) and also well attested in GRAK. In Russian, they seemingly do occur, but seem to be rather limited in use, except for three forms: *ma* ← *mama* ‘mum’, *pa* ← *papa* ‘daddy’, *ba* ← *baba* for a “full” *babuška* ‘granny’.¹⁸ In anticipation of the analysis to follow, it has to be underlined (i) that there are less than 200 instances of such truncated forms in the corpus and (ii) that these tokens are made up of very few types: more than 150 for *ma*, 15 for *pa*, and 22 for *ba*, leaving 4 other forms. (A possible Ukr. *ta* for a TRADVOC *tatu* of the nominative *tato* ‘daddy’ is not attested in the corpus.) The four other forms are of no interest to our aims, by default due to their low quantity. Neither are the former three of interest, because they are common to Ukrainian and Russian.¹⁹ We will not consider this group IV in the analyses to follow, with one exception.

The question arises whether the three remaining analyzable types of expressions of vocative function can shed light on the extent to which Ukrainian Suržyk is shaped by Ukrainian or Russian. The answer can be an unequivocal yes only for instances of the type TRADVOC: this type can definitely be classified as Ukrainian. The two other types are present in Ukrainian and Russian. However, there are clear differences of degree or typicality: We interpret NOMVOC as “more” Russian, due to the fact that it is the normative way to express the vocative function in Standard Russian and very widespread in all other varieties of Russian. The same holds for NEOVOC, although it occurs only in colloquial varieties of Russian.

4.1.2 Analysis

As has been outlined in the introduction to this special issue (Hentschel et al., forthcoming; for more details cf. OCUS-Doc), the corpus consists of two sorts of spoken texts in equal shares: family conversations and guided open interviews: i.e., *family corpus* vs. *interview corpus*. It is not surprising that all but a handful of forms in vocative function are from the former.

The descriptive figures alone already reveal the general tendencies in expressing the vocative function, cf. Table 1.

Table 1 *Frequencies of the three types of vocatives*

Type	N	%
TRADVOC	94	3.8
NEOVOC	392	15.7

¹⁷ These nouns have neuter gender and do not go back to nouns of the mentioned *a*-declension. The decisive conditions are obviously (i) the ending /-a/ in the nominative, here however in the plural, and (ii) the fact that the resulting vocative form without an ending, i.e., the stem, is different from the nominative singular. Masculine nouns with a nominative plural on /-a/, like *professor*, *doktor*, do not exhibit a vocative plural form with a “zero-ending”.

¹⁸ Already Jakobson (1962) stated that the phoneme structure of nursery names for mother and father (and grandmother) are very similar in many diverse languages: so-called babble words produced by babbling babies, sometimes involving multiple repetitions of similar or identical syllables, which are often lexicalized in “adult” language with two syllables. This means that the forms *ma*, *ba*, *pa* are not necessarily truncations in the same sense as *xlo* definitely is.

¹⁹ Note that no instances of a NEOVOC *tat* ← ukr. *tato* ‘daddy’ are witnessed in the corpus. The nominative, however, is on /-o/, not on /-a/. The latter, *tato*, in its full form is rarely witnessed in the corpus compared with Russ. *papa* (the relation is 1 to 7; cf. Hentschel, 2024). The latter, *papa* is not attested in a TRADVOC form: **papo*. See below.

NOMVOC	1,995	79.9
Total	2,481	100

TRADVOC forms, which are clearly Ukrainian, with their share of roughly 4%, obviously play an extremely marginal role. Eighty percent of all forms are NOMVOC, 16% are NEOVOC. The quantitative relation between TRADVOC and NOMVOC in Table 1 corresponds broadly to the observation made by del Gaudio (2010, p. 106), though he does not mention the NEOVOC. With regard to the frequency of the latter, one has to note that this form is possible only for a minority of nouns or names (see above), whereas NOMVOC and (in Ukrainian) TRADVOC are not formally restricted. As previously mentioned, NOMVOC and NEOVOC forms can be considered typical for Russian, but they are also used in non-standard varieties of Ukrainian; however, due to a lack of relevant research, we do not know the extent to which this is the case. However, the figures for vocatives underline observations on regularity in relation to various other phenomena in different subparts of linguistic structures (cf. Hentschel & Palinska, 2022): If in a set of functionally equivalent Ukrainian and Russian expressions, a Russian variant coincides with a Ukrainian non-standard variant, which in addition has a fairly broad distribution in subvarieties of Ukrainian (be they regional or social dialects), then forms corresponding to the Russian and non-Standard Ukrainian expression dominate in Suržyk, and not the normative Ukrainian variant also contained in the set.

Del Gaudio (2010, p. 106) states, furthermore, that young people (students) perceive the TRADVOC as a marker of high formal speech. In Table 2, we therefore test a very frequent lexical item, which is clearly a non-formal one: Ukr. / Russ. *mama* as the affectionate expression of ‘mother’, in contrast to the formal equivalents Ukr. *maty* and Russ. *mat’*. In Table 2—as an exception—we include SYLTrVOC forms as well, as in our corpus they are very frequent for *mama*, but highly marginal for other lexemes (see above).

Table 2 Frequencies of the four types of vocatives for Ukr. / Russ. *mama*

Type	N	%	Form
TRADVOC	16	3.5	<i>mamo</i>
NEOVOC	137	29.8	<i>mam</i>
NOMVOC	148	32.2	<i>mama</i>
SYLTrVOC	158	34.4	<i>ma</i>
Total	459	100	

The two shortened variants of the forms (*mam*, *ma*) clearly dominate: together they have a share of almost two thirds. The other third is made up by NOMVOC forms, again leaving the TRADVOC forms in a marginal position. However, these instances of *mamo* can hardly be described as formally high style. The infrequent occurrences of TRADVOC forms are, to a very high degree, limited to either personal names (about one third) or to kinship terms (one half). Only one out of ten instances consists of a generic noun, almost all of which are affectionate addresses to people close to the speaker: *sestro* ← *sestra* ‘sister’, *dytyno* ← *dytyna* ‘child’, not necessarily to a child, often to a clearly younger person, *panjanko* ← *panjanka* ‘young lady’, *bratyku* ← *bratyk* ‘brother; diminutive’, *holube* ← ‘my dear, lit.: pigeon’, etc. Thus, as rarely as it occurs in Suržyk, the TRADVOC is mainly motivated by affectionate speech. The only high-style uses in the corpus are invocations of the country: *Ukrajino!* All of these names and nouns in TRADVOC (can) occur in NOMVOC as well.

With regard to formal speech, an observation made by Shvedova and Lukashevskyj (2025) is of interest. They studied transcripts of sessions of the Ukrainian Parliament,²⁰ which can only conditionally be considered examples of spontaneous spoken language, and identified a tendency toward a more frequent use of TRADVOC forms in the 1990s and early 2000s, in later years followed by a subsequent decline compared to nominative forms used in the same function. We would interpret this up and (then) down tendency to mean that TRADVOC forms, as clearly specific forms of Ukrainian, were an ideal marker of Ukrainian (non-Russian) identity in the young independent Ukraine, especially in contexts where this is highly useful: in a public political setting such

²⁰ The authors conducted the study based on the UD_Ukrainian_ParlaMint subcorpus (cf. Lukashevskyi et al., 2025). For this corpus, a model was trained to more accurately identify the vocative as a dependency relation. However, text fragments with 100% vocative usage were identified in the data and assumed to reflect editorial normalization; these fragments were therefore excluded from the analysis (Shvedova & Lukashevskyi, 2025, p. 107). Nevertheless, this raises doubts about the reliability of the remaining data on vocatives as well.

as the parliament. Later, this fades away to a certain degree, when the rather neutral NOMVOC forms again gain in frequency. After February 2022, the probability of an increase in the use of TRADVOC forms as overt markers of divergence from Russian may have increased again.

This all suggests that instead of a stylistic contrast between informal NOMVOC and formal TRADVOC the decisive point could be that today the latter is generally outdated in spoken Ukrainian, apart from its marginal role in Suržyk. It can easily function either as an overt marker of Ukrainian identity (or “non-Russianness”) in more or less formal public speech or as a marker of closeness in informal, if not affectionate speech, mainly with personal names of persons emotionally close to the speaker or with kinship expressions.

Further insights emerge when we differentiate the frequencies of vocative types by the stem types Ukrainian, Russian, and common (cf. Hentschel forthcoming b for details on delimiting common stems from Ukrainian and Russian ones). Table 3 presents the proportions:

Table 3 *Types of vocatives and stem types*

stem type / voc type	N			%		
	Ukrainian	Russian	common	Ukrainian	Russian	Common
TRADVOC	33	0	61	24.6	0.0	3.1
NOMVOC	99	342	1,540	73.9	90.0	78.9
NEOVOC	2	38	352	1.5	10.0	18.0
Total	134	380	1,953	100	100	100

First, it should be noted that, in general, the majority of vocatives, irrespective of their type, are based on common stems, shared between Ukrainian and Russian. This is due to the fact that there is a vast overlap between Ukrainian and Russian in personal names and kinship terms.

The second point is that TRADVOC forms are only witnessed when they are based on Ukrainian or common stems, such as Ukr. *dytyno* ‘child’, *diūčjyno* ‘girl’, *matinko* ‘mum’, *Nadije Ivaniúno* (the typical combination of prename and father’s name), or *Kateryno* (prename) based on Ukrainian stems, and *holube* ‘buddy [literally: pigeon]’, *djat’ku* ‘uncle’, *mamo* ‘mum’, *kume* ‘godfather’, *Nazare* (prename; the latter all on common stems). Common stems (lexemes) are most probably perceived by some speakers of Suržyk as Ukrainian, by others, who are more embedded in Russian-speaking social contexts, as Russian, as is evidenced by the papers by Hentschel (forthcoming a) on inflectional morphology and Zeller (forthcoming) on phonic structures. There is not one instance in the corpus in which a TRADVOC form is built on a Russian stem, although structurally they could have been, e.g., Russ. *papa* → **papo*, *devuška* → **devuško* (literally ‘girl’, used as a, not necessarily, affectionate form of address not only for girls but also for not “really” young ladies), *lastoč’ka* → **lastoč’ko* (literally ‘swallow’, widely used for an affectionate address for girls, young ladies), *Vladik* → **Ŭladiku* (personal name).

The latter observation contrasts somewhat with the general findings on inflectional morphology in Suržyk (cf. Hentschel forthcoming b). Most inflectional constellations of Russian stems in nouns, adjectives, and verbs (e.g., CNG for the former two, TPN or TGN for the latter) show a very strong or strong tendency to be adapted to Ukrainian inflectional patterns, i.e., to be combined with Ukrainian endings. Few inflectional constellations do so to a weaker or much weaker degree, such that these cases can be plausibilized in preference theoretical terms (Hentschel forthcoming b). The absolute lack of Russian stems with a Ukrainian vocative ending thus hints at a development in which TRADVOC inflectional forms have widely lost their productivity in Ukrainian itself and gradually turn into emblematic forms, in more or less the same way as the form Russ. *otče* (= Ukr. *otče*) is used to address a priest. To be precise, native speakers of Ukrainian know how to build TRADVOC forms, but they don’t (notwithstanding the emblematic forms), unless a formal communicative situation demands conscious Ukrainian speech.

The third and final point to be made in the discussion of different stem types is to underline that NEOVOC forms are extremely rare with specifically Ukrainian stems. Only one form is attested: twice coll. Ukr. *doc’* ← *docja* ‘daughter’ (cf. coll. Russ. *doča*). This gives strong support to Taranenko’s hypothesis of the Russian origin of NEOVOC forms (Taranenko, 2024, pp. 97–104).

In addition to the general tendencies in the differences between the three vocative types outlined above, regional differences in their distribution also exist. We follow the regional differentiation proposed by Hentschel and Taranenکو (2021) on the basis of the presence of the three “codes” Ukrainian, Russian, and Suržyk in respondents’ everyday lives. The data were collected in two large-scale sociodemographic surveys and later partly incorporated into OCUS (see also the map in Hentschel et al., forthcoming). Table 4 lists the names of the oblasts where data were collected and Table 5 shows the regional differences:

Table 4 *Regions and oblasts*

	Oblasts
Region I	Xmel’nyč’kyj, Čerkasy
Region II	Vinnycja, Kyiv, Kirovohrads’ka, Žytomyr, Čerkasy, Černihiv
Region III	Sumy, Dnipropetrovs’ka, Odesa
Region IV	Xerson, Xarkiv
Region V	Mykolaiv

For the Regions I and II, a more or less obvious quantitative dominance of Ukrainian over the other two codes is characteristic, most clearly in Region I. In Region III it is Suržyk that dominates, in IV and V, it is Russian, most clearly in V, with the other codes exhibiting a reduced presence (see also the map in Hentschel et al., forthcoming). By and large, Regions I and II represent the more western or more central (not peripheral, border) parts of Ukraine; Regions III, IV and V represent the eastern and southern parts of the country.

Table 5 *The three vocative types in the five regions*

Tab. 5a	N					
	Region I	Region II	Region III	Region IV	Region V	Total
TRADVOC	59	24	4	7	0	94
NEOVOC	41	63	103	87	98	392
NOMVOC	439	436	428	430	262	1,995
Total	539	523	535	524	360	2,481
Tab. 5b	in %					
	Region I	Region II	Region III	Region IV	Region V	
TRADVOC	10.9	4.6	0.7	1.3	0.0	
NEOVOC	7.6	12.0	19.3	16.6	27.2	
NOMVOC	81.4	83.4	80.0	82.1	72.8	
Total	100	100	100	100	100	

The figures for the vocative types are compatible with the regional classification in Hentschel and Taranenکو (2021): the typically Ukrainian TRADVOC shows the highest values in Regions I and II, especially in I. In both regions the retreat of TRADVOC is delayed. In Regions III to V it virtually plays no role at all, except for very few exceptions. The values for the “more Russian-like” NEOVOC exhibit the opposite tendency: The weakest values are to be found in Regions I and II, especially in I, but NEOVOC is much stronger in Regions III to V, with the strongest value in V. The NOMVOC is relatively stable in all regions. (We cannot explain the rather smooth decline from a little more than 80% in Regions I to IV to a little more than 70% in Region V.)

Last but not least, there is a further sociobiographical feature that obviously has an influence on the frequency of the three variants of vocative expressions: age.

Table 6 *Vocative types and the age of respondents*

	N			
Age group / Voc type	< 29 years	29–49	> 49	Total
TRADVOC	9	47	38	94

NEOVOC	173	168	40	381
NOMVOC	505	922	516	1,943
Total	687	1,137	594	2,418 ^a
	in %			
TRADVOC	1.3	4.1	6.4	
NEOVOC	25.2	14.8	6.7	
NOMVOC	73.5	81.1	86.9	
Total	100	100	100	

^a Please note that the total number here is 66 instances lower than in Tab. 5, because these instances stem from respondents whose age is not known.

Two observations are most relevant: (i) TRADVOC forms, though generally rare, are least frequent with the youngest respondents (under 29 years). The value for the oldest respondents (50 years and older) is more than four times higher. (ii) Still clearer is the decrease in the frequency of NeoVoc forms from the group of the youngest respondents to the oldest. On the other hand, the values for NOMVOC forms, which in all three age groups shows the highest values by far, are relatively stable, with a slight decline in the younger age groups.

Region I and, to a lesser extent, Region II are the only regions still demonstrating a considerable share of TRADVOC forms, whereas in the other three regions the TRADVOC has almost disappeared. Thus, the role of age, especially in Region I, is of interest:

Table 7 *The three vocative types and the age of respondents in Region I*

Region I	N			
age group / Voc type	< 29 years	29–49	> 49	Total
TRADVOC	3	29	27	59
NEOVOC	26	12	3	41
NOMVOC	100	237	107	444
Total	129	278	137	544
	in %			
TRADVOC	2.3	10.4	19.7	
NEOVOC	20.2	4.3	2.2	
NOMVOC	77.5	85.3	78.3	
Total	100	100	100	

It is very clear for Region I that TRADVOC forms have the widest distribution among older age groups, especially the oldest respondents, whereas for these age groups NEOVOC forms play only a marginal role.

When it comes to Region II, it suffices to state that the differences between the age groups for the usage of TRADVOC forms vanish: they are peripheral for all three age groups. All other figures agree with the general distribution as illustrated in Tab. 6.

At first sight, the age-based results for NEOVOC in the five regions (especially in Region I) contradict the hypothesis that the influence of Russian at least plays a supporting role in the spread of NEOVOC forms and thus conflicts with the results in Tab. 5. NEOVOC forms are most frequent in the youngest age group, i.e., the generation that widely grew up in independent Ukraine, when political and social pressure in favor of Russian was lowest. The decisive point seems to lie in the “nature” of NEOVOC as a marker of informal, casual, affectionate speech in Russian. The spread of NEOVOC in the youngest age group may have been reinforced by their much greater usage of social media and internet communication in general, where informal and casual speech dominates, both in Ukrainian and Russian, cf. as well Gradus (2026). Due to the general and doubtlessly

long ongoing retreat of TRADVOC in (spoken) Ukrainian, which is definitely strongest in the young generation, NEOVOC fulfills this stylistic discourse-functional gap more effectively than NOMVOC. And younger people, without doubt, are more inclined to express themselves in a casual way, which is different from the way older people express themselves, even in a low-variety like Suržyk. This interpretation is, of course, a hypothesis.

One final factor that might hypothetically influence the choice of forms with vocative function is the respondent's exposure to Russian in everyday life. Hentschel (forthcoming a) introduced this complex, gradient parameter and demonstrated its effect on the tendency to adapt Russian stems to Ukrainian inflectional patterns: the greater the exposure to Russian, the greater the tendency to retain Russian inflectional endings rather than replacing them with Ukrainian ones.

The usefulness of this parameter for analyzing the distribution of the three types of vocative expressions seems at first glance to be rather limited, since of the three types, only the very infrequent TRADVOC can be regarded as unequivocally and exclusively Ukrainian. By contrast, the two standard realizations of the vocative function in Russian, NEOVOC and NOMVOC, have—if the literature reviewed above is correct—already become firmly established in Ukrainian, quite apart from their use in Suržyk. Nevertheless, respondents' exposure to Russian does play a role, albeit a rather limited one. Table 8 presents the results of the quantitative analysis for this parameter.

Table 8 *Exposure to Russian and the three vocative types*

	Vocative type N			N all	Vocative type %		
Exposure to Russian	NEOVOC	NOMVOC	TRADVOC		NEOVOC	NOMVOC	TRADVOC
Weak	283	1,617	87	1,987	14.2	81.4	4.4
Medium	85	314	6	405	21.0	77.5	1.5
Strong	24	64	1	89	27.0	71.9	1.1

In general, the quantitative distribution of the three vocative types differs only slightly across the groups with different degrees of exposure to Russian. In all three groups, the clearly dominant type is NOMVOC, accounting for approximately three quarters of all occurrences. Nevertheless, its share shows a gradual and remarkably consistent decline from the group with weak exposure to Russian to that with strong exposure. Thus, the high frequency of NOMVOC is clearly not reinforced by exposure to Russian; if anything, the opposite is true. This provides further evidence that NOMVOC occupies a strong position in Ukrainian itself.

The second most frequent type in all three groups is NEOVOC, with proportions ranging from 14% to 27%. Here, however, the pattern is the reverse: its share almost doubles from the group with weak exposure to Russian to the group with strong exposure. In other words, the Suržyk data suggest that NEOVOC may be continuing to expand in spoken Ukrainian, even independently of further processes of Suržykization, simply as a consequence of continued exposure to Russian in everyday life.

Finally, the absolute figures in Table 8 provide convincing evidence that TRADVOC occurs almost exclusively among speakers with only weak exposure to Russian.

This all leads to the conclusion for vocatives that NOMVOC is the unmarked expression in Suržyk (and most probably in spoken, more or less informal varieties of Ukrainian in general), whereas TRADVOC is in a process of strong decline (again, not only in Suržyk) or has even widely lost its productivity in unconscious forms of Ukrainian speech. In addition to emblematic usage it remains either a selective residue especially for older people in affectionate speech or a means to overtly signal “Ukrainianness” as opposed to “Russianness” in conscious speech, the latter most probably being of lesser importance in conversations in Suržyk. The use of NEOVOC obviously increases, especially when Russian has played an important role in people's everyday life.

4.2 Non-vocative discourse-structuring elements

4.2.1 Introductory remarks

In general, the following analysis builds on Hentschel (2024), which was devoted primarily to lexical categories with “full” semantic content, such as nouns, verbs, and adjectives. The individual lexemes constituting a hyperlexematic pair—Ukrainian on the one hand and Russian on the other—could be regarded as Ukrainian-Russian translation equivalents. In the case of discourse markers, however, this concept of contrasting Ukrainian and Russian units reaches its limits. The fundamental reason is that many of these elements have undergone semantic bleaching to a greater or lesser extent. Consequently, in the case of discourse markers, it is more appropriate to speak of functional rather than translational equivalence.

Functional equivalence, however, also presents a problem. This can be illustrated by the example of hesitation markers, or fillers occurring during pauses in speech, as a subgroup of discourse markers. Hesitation markers belong to a set of expressions that Russian and Ukrainian linguists (with a rather normative approach in the sense of the *culture of speech*, Ukr. *kul'tura movlennja*, Russ. *kul'tura reči*) call *parasitic words* (Ukr. *slova-parazyty*, Russ. *slova-parazyty*). To the best of our knowledge, no such metaphorically emotional, pejorative term based on a metaphor of parasitism has been established in the linguistic tradition in Western languages. A more appropriate term seems to be *disfluency marker* (cf. the discussion in Crible, 2018). We will draw our examples for hesitation markers from Russian, as this area has been studied more extensively. Apart from the typical non-lexical elements such as *é(-è)*, *m(m)*, *ém*, and the like, which, as noted above, are excluded from the present analysis, the following items, among others, may be assigned to this class: *ëto*, *ëto samoe*, *ëtot*, *nu*, *vot*, *značit*, *kak by*, and *tipa*. In some contexts, where these items merely serve to fill a brief pause in speech (or, from the perspective of utterance analysis, where no other motivation for their use can be identified), variation among them is likely to be largely free. In other contexts, however, this is probably not the case. Correspondingly, each of these elements is generally associated with several functions. For example, *nu* may function as a response marker, a turn-management marker, or a hesitation marker, while *ëto* may serve as a hesitation marker, a search marker, or a repair marker (cf. Bogdanova-Beglarian et al., 2018; Burgsteden, 2023). Thus, the two elements overlap in their hesitative function and may be regarded as interchangeable in this use, whereas they are (at least generally) not interchangeable in their respective non-shared functions. Even within a single language, therefore, functional equivalence proves to be a complex matter; from a contrastive perspective, this complexity naturally increases further.

Bogdanova-Beglarian et al. (2018) distinguish 20 subfunctions. The analysis proposed here cannot achieve such a level of differentiation. In other words, the following analyses do not claim to provide an exhaustive account of discourse-organizing expressions in Suržyk. Such a task would require a comprehensive, presumably monograph-length study. Instead, the present analysis is restricted to those expressions that can be identified relatively easily and semi-automatically in syntactically non-integrated positions on the basis of their form, their (at least approximate) grammatical annotation by word class, and the punctuation or other formal devices to mark pauses used in OCUS. A further requirement is that they occur as functionally equivalent Ukrainian-Russian pairs (or triples, etc.) with sufficient token frequency in OCUS to permit an examination of the potential influence of sociobiographical factors such as regional origin, age, or education.

For the present study, 16 Ukrainian-Russian “hyperlexematic” items, that is, Ukrainian-Russian pairs of discourse-structuring elements, were selected (see Table 9 below). (In the notation of hyperlexemes adopted here, the Ukrainian member always appears to the left of the equals sign and the Russian member to the right.) The number of forms included in the analysis, however, is not simply twice the number of hyperlexematic items. Instead, a total of 70 forms were taken into account. For most hyperlexematic items, more than one form was included on both the Ukrainian and the Russian side, provided that these forms shared the general function of the corresponding lexematic (lemmatic) item. Altogether, the analysis is based on more than 17,000 tokens. The minimum frequency required for inclusion was 150 tokens,²¹ whereas the maximum observed frequency was approximately 6,400 tokens.

4.2.2 A general overview of the data

²¹ For example, *do reči* = *kstati* ‘by the way’, *po-perše* = *vo-pervyx* ‘first(ly)’, *po-druhe* = *vo-vtoryx* ‘second(ly)’ and several others were below that limit.

Table 9 lists the items analyzed. The rows follow the relative token frequency (in %) of the Ukrainian variant of the hyperlexeme illustrated in the rightmost column, in descending order. The leftmost column thus mirrors the rank of that order. The difference to 100% is the share of the Russian variant.

Table 9 *The discourse markers tested²² and the relative frequencies of the variants*

No./rank	hyperlexeme	Elements (dm-lexemes / forms)	N all	N ukr	% ukr
1	(ne)xaj = pust', puskaj	xaj (734), nexaj (105) = pust' (17), puskaj (29)	885	839	94.8
2	(po)dyvytysja = (po)smotret' ^a	dyvy (86), dyvys' (185), dyvit'sja (53), podyvy (5), podyvys' (17), podyvit'sja (5) = smotri (20), smotrite (3)	374	351	93.9
3	čuty = slyšat' ^a	čuješ (143), čujete (3) = slyšiš' (6), slyšite (1), slyš' (3)	156	146	93.6
4	(po)sluxaty = (po)slušat' ^a	sluxaj (121), posluxaj (9), sluxajte (11), posluxajte (1) = slušaj (20), poslušaj (2), slušajte (3)	167	142	85.0
5	os', ot, osjo = vot	os' (377), ot (3,205), oto (807), osjo (305), osjo-ho (5) = vot (1,706), vo (1)	6,406	4,699	73.4
6	cej = ètot ^b	cej (293), ocej (8) = ètot (111)	412	301	73.1
7	ce = èto ^c	ce (176), ce-vo (8), oce (97), oce-vo (2) = èto (200)	483	283	58.6
8	jak by = kak by	jak by (399) = kak by (503)	902	399	44.2
9	napryklad = naprimer	napryklad (431), do prykladu (1) = naprimer (513), k primeru (74)	1,019	432	42.4
10	napevno, mabut' = naverno ^d	mabut' (262), napevno (29), napevne (1) = naverno (995), naverno (23), nebos' (1), verojatno (1)	1,312	292	22.3
11	dobre = xorošo	dobre (90), dobren'ko (2), harazd (1) = xorošo (337), ladno (125)	555	93	16.8
12	rozumity = ponimat'	rozumiješ (12), rozumijete (53) = ponimaeš' (306), ponimaete (160)	531	65	12.2
13	tobto = to est'	tobto (100) = to est' (1,217)	1,317	100	7.6
14	zrozumilo = ponjatno ^e	zrozumilo (13) = ponjatno (193), ponjatnen'ko (5)	211	13	6.2
15	zvyčajno = konečno ^e	zvyčajno (70), zvisno (25) = konečno (1,588)	1,683	95	5.6
16	typu = tipa ^f	typu (29) = tipa (834)	863	29	3.4

^a All uses as “full verbs” have been ignored.

^b All uses as demonstratives have been ignored.

^c All uses as demonstratives and conjunctions have been ignored.

^d For Ukr. *napevno*, *napevne* and Russ. *naverno*, *naverno* only uses in the sense of ‘probably’ have been included, i.e., as synonyms of the other expressions here.

²² The hyperlexemes postulated represent the bilingual default relations as they are codified in major Ukrainian-Russian and Russian-Ukrainian dictionaries (Il'in & Dorošenko, 1965; Ižakevič et al, 2004; Žajvoronok, 2003).

^e All uses as adverbs have been ignored.

^f All uses as “full nouns” have been ignored.

Before going into the analysis, the usage of these elements is illustrated by two examples for each hyperlexeme. All illustrations contain one utterance from the corpus [a] with the Ukrainian variant of the hyperlexeme and [b] with the Russian one, always preceded by the corresponding number from Table 9:²³

- [1a] *Nexaj*, može, jakos' protjanem.
'Well, maybe we'll somehow get by.'
- [1b] Nu, karočje, ja napysala kontrol'nu na sto baloŭ, *puskaj*.
'Now, in short, I got a hundred on the test, *that's it*.'
- [2a] Oj, *dyyys'*, my bahato hovoryly s mojimy druž'jamy.
'Oh, *look*, we talked a lot with my friends.'
- [2b] U nas, palučajícca, *smatrití*, eta sélo hraničit s moldavijij.
'In our area, it turns out, *look*, this village borders on Moldova.'
- [3a] Tolku s toho, šo vona pokazuje, *čjuješ?*
'What's the point of that, that she's showing [it], [*do you*] *hear?*'
- [3b] Nada bulo, *slyšyš*, pryjty.
'It was necessary, [*do you*] *hear*, to come.'
- [4a] Hupky krasivi, *sluxaj!*
'The lips are beautiful, *listen!*'
- [4b] *Slušaj*, možna dva dnja ubyrat' kvartiru.
'*Listen*, it could take two days to clean the apartment.'
- [5a] Ot, tak i todi hovoryly.
'So [you know], that's how they spoke then.'
- [5b] Vona z zapadnoj pryjixala, vot.
'She came from the West [Ukraine], so [you know].'
- [6a] Tak šo, *cej*, je šo ũspomnyt'.
'So, *this* [well], there is something to remember.'
- [6b] Ty ũ t'oti, *étot*, Ani rozkazuvala.
'You at Aunt, *this* [= what was her name?], Anja's [place] told this.'
- [7a] Babuška ũ nas *ce* (...) rano pomerla.
'Our grandmother, *well*, (...) died early.'
- [7b] Vin, *éta*, pašol apjat' v bal'nicu.
'He, *well*, went to the hospital again.'
- [8a] Rodylys' v Ukrajini, *jak by* ...
'[We] were born in Ukraine, *as if* [like]...'

²³ Note that in the English paraphrases the discourse markers are, in most cases, noted as a literal translation. These translations are in some cases without doubt pragmatically rather not equivalent to an English utterance. In these cases, a (hopefully) better equivalent is proposed in square brackets.

- [8b] No, *kak by*, raditíl'j my nazyvaly úsígda na ukrains'kom.
'But, *as if* [like], we always spoke to our parents in Ukrainian.'
- [9a] Ja ljublju, *napryklat*, vyšyvat'.
'I like, *for example*, to embroider.'
- [9b] Nu, ot, *naprimjer*, v institutje...
'Well, so, *for example*, at the institute...'
- [10a] Kraščě, *napeŭno*, tam, de dva jazyka.
'In short, *probably*, there, where there are two languages.'
- [10b] Ce vono tak, *navjerno*, polučjylos'.
'So that's how it turned out, *probably*.'
- [11a] *Dobre*, davaj ja tobi peredzvonju.
'Good [Okay], let me call you back.'
- [11b] *Xarašo*, my postarajemos'.
'Good [Okay], we'll try'
- [12a] My, *rozumiješ*, nikoly ne spišymo.
'We, *you understand*, never hurry'
- [12b] Ėta tak, *ponimajíš*, ŭ narodi kažut'.
'That's what, *you understand*, people say.'
- [13a] Nu, *tobto*, jak ob"jasnyt'...
'Well, *that is*, how to explain ...'
- [13b] I, *to jest*', ŭše bude rozvyvac'eja.
'And, *that is*, everything will develop.'
- [14a] *Zrozumilo*, bo ja ne chočju borščja.
'Understood [clear], because I don't want borsch.'
- [14b] Nu, tobi, *ponjatno*, vono ně nado.
'Well, you, *understandably* [clearly], don't need that.'
- [15a] Nu i ditej bil'sě stalo, *zvyčajno*.
'Well, and there are more children now, *of course*.'
- [15b] *Kanješno*, chlib zaras pečjut' duže dobryj.
'Of course, bread is baked very well now.'
- [16a] *Typu*, vízdje možu hovoryty.
'Like, I can speak everywhere.'
- [16b] Vin, *tipa*, imjeje pravo.
'He, *like*, has the right [to do s.th.].'

4.2.3 Interpretation of the quantitative results

Readers who are less familiar with quantitative methods²⁴ may wonder whether it is possible to draw any general conclusions from “only” 16 hyperlexemes of discourse markers, represented by roughly 17,000 tokens. The same question could be raised with respect to Hentschel (2024), who examined just over 100 lexemes (predominantly nouns, adjectives, verbs, adverbs, as noted above) which were, however, represented by approximately 100,000 tokens—almost one fifth of the relevant corpus material for that study. In our view, the answer to this question is positive. The crucial question, however, is not whether generalizations are possible, but rather what kind of generalizations can be made and to what extent they are legitimate.

In this respect, two general observations should be made concerning the discourse markers examined here. (i) To date, no study has systematically investigated discourse markers in Suržyk; at least there are no corpus-based studies. It is therefore reasonable to assume that, in principle, Suržyk employs the same discourse markers as Ukrainian (including its local dialects) and Russian. Whether these markers are used differently in Suržyk in one respect or another to their use in the two source languages (especially in their codified standard varieties) cannot be assessed within the scope of the present study. (ii) As noted above, the analysis includes 70 Ukrainian or Russian expressions, represented by approximately 17,000 tokens, that function as discourse markers in either Ukrainian or Russian and also occur with this function in the underlying OCUS corpus. Their specific discourse functions are not distinguished here, although, as discussed above, many of these forms are likely to fulfill several different functions. Ukrainian and Russian, of course, also share a number of discourse markers, such as the already mentioned *nu*, which occurs more than 19,000 times in the corpus. Such shared items are naturally excluded from the analysis.²⁵ Given these figures,²⁶ it seems reasonable to assume that the great majority of discourse markers occurring in OCUS have been covered. Considering this background, the data presented in Table 9 allow the following general observations.

It is immediately apparent that the discourse-marker forms attested in OCUS display the same overall pattern of preference for either the Ukrainian or the Russian variant of the corresponding hyperlexeme as Hentschel (2024) found for the lexicon as a whole.²⁷ Some hyperlexemes are realized almost exclusively by Ukrainian forms, whereas others are realized almost exclusively by Russian ones. If a threshold of 85% (or more) versus 15% (or less) is adopted, hyperlexemes 1–4 belong to the first group, showing a clear predominance of Ukrainian realizations. Hyperlexemes 12–16 (and, approximately, 11) belong to the second group, characterized by a clear predominance of Russian realizations. We will refer to these two groups as “Ukrainian-dominated” and “Russian-dominated” hyperlexemes, respectively.

These dominant patterns suggest that one variant (mostly one form)—whether Ukrainian or Russian—has become largely stabilized as the conventional form in Suržyk. The non-dominant variants may therefore be interpreted as instances of sporadic, spontaneous influence from either Ukrainian or Russian. According to Auer (1999), a fused lect, that is, a mixed variety containing a large number of quasi-obligatory elements from one or the other source language (with potentially different preferences across phonological, grammatical, or lexical variables), may nevertheless exhibit occasional departures from the established form in favor of the corresponding variant from the other source language, as long as both source languages continue to be used within the speech community. This has, of course, been the case in Ukraine until recently, although the use of

²⁴ Diffuse scepticism toward quantitative analyses is often expressed by scholars who do not themselves work with quantitative methods. Where they engage in empirical analyses of linguistic data at all, especially in the study of variation in non-standardized varieties, they typically rely on their own—often extensive—collections of examples, which are evaluated informally. To what extent intuition, prior assumptions, or other forms of background knowledge have influenced the selection of the data is rarely made explicit. Such analyses are generally based on chance observations, the discussion of particularly illustrative examples, and holistic impressions of the patterns under investigation. In Suržyk research, this applies even to the most comprehensive and, in many respects, highly valuable account by del Gaudio (2011), as well as to the work of Šumarova (2014). Should one assume that analyses based primarily on informal, sometimes impressionistic evidence provide a sounder basis for generalization?

²⁵ Other examples where there is no difference between the Ukrainian and Russian expressions or the difference is minimal include *korotše* = *koroče* ‘shortly (saying)’, *dopustym(o)* = *dopustim* ‘let’s guess’, *pravyl’no* = *pravil’no* ‘right’ etc.

²⁶ The discourse markers included in the present analysis, together with *nu*, account for approximately 5% of all tokens in OCUS. For comparison, pronouns—the most frequent word class in the corpus—account for approximately 17%.

²⁷ As a matter of fact, Hentschel (2024) included three of the hyperlexemes from Table 8 above. However, he did not differentiate between their use in syntactically isolated positions as discourse markers and their use as semantically “full” lexemes. Be that as it may, the results of both analyses are highly compatible.

Russian appears to have declined gradually over the past two or three decades and much more sharply since February 2022.

Between these two groups, Table 9 reveals a third group, characterized by a more or less substantial degree of variation between the Ukrainian and Russian variants. This is particularly evident for hyperlexemes 7, 8, and 9, for which the Ukrainian and Russian variants occur in roughly balanced proportions. (These are highlighted in Table 9 by a darker shade of blue.) Considerable variation is also found for hyperlexemes 5 and 6 (highlighted in a lighter shade of blue), where, despite the clear predominance of the Ukrainian variant, its share remains below 75%. In the following section, these five hyperlexemes will be examined in greater detail in order to identify the factors underlying this variation, taking their potential interrelationships into account.

Taken together, the results show that the forms functioning as discourse markers in Surżyk are derived from both Ukrainian and Russian in relatively balanced proportions. This general pattern manifests itself in two ways. First, there is a clear tendency for either the Ukrainian or the Russian variants of individual hyperlexemes to become established as the conventional form in Surżyk. Second, for some hyperlexemes, the Ukrainian and Russian variants occur in relatively balanced proportions. These cases will be examined in greater detail in the following section, with particular attention to their dependence on sociobiographical factors and their relationship to other linguistic preferences. By contrast, the present study cannot explain why, in some hyperlexemes, the Ukrainian variants have become established, whereas in others it is the Russian variants that have prevailed. The reasons are likely to be highly specific to the individual hyperlexemes and, if they can be identified at all, would require detailed case-by-case analyses.

4.2.4 On the five instances of variation

For the five hyperlexemes [5] through [9] with a considerable degree of variation between Ukrainian and Russian expressions, the factors underlying that variation are now examined. Mixed-effects models allow for the inclusion of random and fixed effects (non-terminologically referred to here as *factors*; cf. Döring & Bortz, 2016, pp. 897, 915). The aim is to identify the latter. Random effects, which may also influence the results, must be specified in advance. In the present study, this is the individual respondents, that is, speakers with their personal preferences and differing propensities to use particular discourse markers. The five hyperlexematic discourse markers themselves, which may initially show a tendency toward Ukrainian or Russian realizations before stabilizing on one option, are analyzed separately, since the reasons why one variant prevails over another may be lexeme- or form-specific and are not limited to intralinguistic factors.

We conduct two analyses for each of the five hyperlexematic discourse markers. First, we examine the influence of sociobiographical and biographical factors. The following variables were included as potential fixed effects: respondent's age, gender, education, their self-reported exposure to Russian (cf. Hentschel forthcoming a), region (following Hentschel & Taranenko, 2021),²⁸ and subcorpus, reflecting the interactional setting (family conversations vs. semi-structured, open interviews—cf. Hentschel et al., forthcoming). These are, in the strict sense, explanatory factors for which a causal relationship with a tendency toward more Ukrainian or more Russian discourse markers is plausible.

In separate analyses, a non-sociobiographical variable was examined, namely a speaker's tendency to adapt Russian lexemes in Surżyk to Ukrainian inflectional patterns. This tendency, identified by Hentschel (forthcoming a) in four clusters of respondents, serves as an indicator of speakers' orientation toward Ukrainian or Russian. This is not an explanatory variable in the strict sense, but rather concerns correlations between linguistic behavior at different linguistic levels. Briefly, the hypothesis is as follows: Speakers who lean more toward Russian at the inflectional-morphological level are also more likely to use Russian discourse markers. Since the two extreme clusters of the four (i.e., I and IV) are represented by only a few speakers and, in some cases, there are correspondingly few instances of hyperlexematic discourse markers—which in some cases makes the statistical models impossible to calculate—we combine the two clusters that lean more toward Ukrainian (I and II) and the two clusters that lean more toward Russian (III, IV). The differences in the respondents' inclination to adopt Russian stems to Ukrainian inflectional patterns depends, as Hentschel (forthcoming a) shows, to a certain degree on the sociobiographical factor of the respondent's exposure to Russian.

²⁸ See above, Table 4.

The analyses tested the contrast between Ukrainian variants on the one hand and Russian variants on the other; the latter were therefore grouped together. Statistical analyses were conducted in R version 4.3.2 (R Core Team 2023), using the *lme4* and *lmerTest* packages (Bates et al., 2015; Kuznetsova et al., 2017).

(A) On the influence of sociobiographical factors

- Hyperlexeme 5 *os', ot, osjo = vot*

Exposure to Russian and age were significant; the higher the exposure to Russian, the less likely the Ukrainian variants are ($\beta = -0.14$, $p < 0.001$). The older the respondents, the more likely the Ukrainian variants are ($\beta = 0.02$, $p = 0.010$). The patterns also vary by region ($\chi^2(4) = 24.09$, $p < 0.001$), even though the results do not exactly match expectations. As expected, the probability of Ukrainian variants is highest in Regions I and II, where according to Hentschel and Taranenko (2021) Ukrainian is much better established as in the other Regions, however unexpectedly in II even higher than in I. „It is lowest in Region III, where Suržyk plays a dominating role in people’s everyday life, even lower than in IV and V, where Russian is dominating (see Appendix). Gender, education, and subcorpus have no significant influence.

- Hyperlexeme 6 *cej = étot*

None of the factors tested were significant.

- Hyperlexeme 7 *ce = éto*

None of the factors tested were significant.

- Hyperlexeme 8 *jak by = kak by*

The Ukrainian variant is more likely in the interview corpus than in the family corpus ($\beta = 8.60$, $p = 0.003$). With greater exposure to Russian, the Ukrainian variant becomes less likely ($\beta = -0.42$, $p = 0.003$). All other factors had no significant influence.

- Hyperlexeme 9 *napryklad = naprimer*

The Ukrainian variant is more likely in the interview corpus than in the family corpus ($\beta = 3.38$, $p = 0.008$). With higher exposure to Russian, the Ukrainian variant becomes less likely ($\beta = -0.18$, $p = 0.030$). Among older respondents, the Ukrainian variant is more likely ($\beta = -0.07$, $p = 0.025$). The remaining factors had no significant influence.

(B) On interdependencies with inflectional preferences

Respondents belonging to the “inflectional clusters” III and IV (with a considerable tendency to realize Russian stems with Russian endings) showed significantly fewer Ukrainian realizations for the following discourse marker hyperlexemes:

- Hyperlexeme 5 *os', ot, osjo = vot* ($\beta = -2.35$, $p < 0.001$)
- Hyperlexeme 8 *jak by = kak by* ($\beta = -11.54$, $p = 0.005$)
- Hyperlexeme 9 *napryklad = naprimer* ($\beta = -3.56$, $p < 0.001$)

The difference was not significant for 6 *cej = étot* and 7 *ce = éto*.

In summary, the following can be stated: Two of the pairs are not influenced by sociobiographical factors, namely 6 and 7, i.e., *cej = étot* and *ce = éto*. Nor do these pairs correlate with the respondents’ behavior at the inflectional-morphological level.

For the remaining three hyperlexematic pairs 5, 8, 9 it holds that:

- i. Higher exposure to Russian is associated with a lower probability of using the Ukrainian variant(s) in all three pairs.
- ii. In two of the three pairs, older respondents use the Ukrainian variant(s) more frequently than younger respondents (5 and 9).
- iii. For two of the three pairs, the Ukrainian variant(s) occur more frequently in the interview corpus than in the family corpus (8 and 9).
- iv. For these three pairs, the probability of using the Ukrainian/Russian variant(s) correlates with behavior at the inflectional-morphological level.
- v. Gender and education do not play any role; region is a factor only in the case of 5.

Explanations for these differences and similarities must be widely left open for detailed studies on each of them, taking into account a greater number of potentially influencing factors than could be done in this overview. The following observation could form the basis for a hypothesis: Hyperlexemes 6 and 7, *cej* = *étot* and *ce* = *éto*, have completely lost their basic function realized in demonstrative pronouns. They simply function as hesitation or disfluency markers, the Ukrainian and Russian expressions occur in free variation. As such they do not exhibit any correlation with the factors tested, neither with sociobiographical nor intralinguistic ones. For the other three hyperlexemes, a function (or in the case of 9, *napryklad* = *naprimer*, even meaning) can be stated positively and thus correlations with other factors (mainly exposure to Russian) can be observed.

5 Summarizing discussion and conclusion

The guiding question for this study was whether syntactically isolated expressions with an organizing function in discourse exhibit a specific inclination to be realized with Russian variants rather than with their functionally equivalent Ukrainian counterparts. The study was motivated by the assumption that linguistic expressions in syntactic isolation are much easier to transfer from a socially dominant language in a society (in our case, for well over a century, Russian) into the speech of speakers of a dominated language (until at least 1990 Ukrainian), even in cases when the latter have a rather weak command of the superstrate.

The figures for vocatives, i.e., for the different forms with vocative function, seem to support the hypothesis of an extraordinarily strong impact of Russian. The types of expression that are either the norm in Standard Russian or a typical and widespread trait of informal, conversational Russian, i.e., the nominative vocative (NOMVOC) and the endless neo-vocative (NEOVOC), absolutely dominate the distribution in Suržyk, together amounting to almost 100% of the forms analyzed. The “clearly Ukrainian” expression of the inherited traditional vocative (TRADVOC) only plays a peripheral role with a share of less than 5%.

Nevertheless, our corpus data—which were collected in two stages about 10 or four years ago—do not allow for an unequivocal conclusion regarding an especially high impact of Russian on vocative expressions in Ukrainian Suržyk. At least this holds for the most frequent NOMVOC with its share of roughly 80%. Already three quarters of a century ago some students of the phenomenon spoke of a general process of the spread of NOMVOC forms at the expense of TRADVOC in Ukrainian itself. Without any doubt, the Russian linguistic conventions of addressing (first and foremost) people had an impact on the corresponding conventions in Ukrainian, especially in Soviet times. However, the overwhelming quantitative dominance of “Russian-like” vocatives cannot be considered an exclusive trait of Suržyk, though here, in the “Suržykization” or hybridization of the speech of Ukrainians, the process of giving up the TRADVOC may have been accelerated or reinforced. The point that the strong “Russian-like” pattern of vocatives in Suržyk is not specific to Suržyk but rather widespread in spoken varieties of Ukrainian itself is supported by the fact that the figures for NOMVOC differ only slightly, not only for the different regions for which data have been collected, but also for respondents of different age groups.

The situation is completely different for TRADVOC on the one hand and NEOVOC on the other. The former is, within tight bounds, better represented in the two western, linguistically Ukrainian-dominated regions differentiated in the analysis, and this rather among older respondents, and does not play any role among younger respondents in the other three, more southern or eastern regions (including Sumy on the northern border with Russia). The distributional characteristics of NEOVOC are diametrically opposed: It is rather common with younger people in the South and the East, but rather uncommon with older people in the West. It is remarkable that the clearly Russian-based NEOVOC is becoming more frequent among younger people—not only to the detriment of Ukrainian TRADVOC, but even to the detriment of the (most probably only superficially) Russian NOMVOC.

These observations can hardly be explained by the process of Suržykization alone; rather, to a certain degree they reflect different social conventions of addressing people in Ukrainian itself that vary along the dimensions

of age and region, notwithstanding the “descriptions” in normative grammars of Ukrainian. Nevertheless, in regard to vocative expressions, Ukrainian Suržyk strongly resembles the patterns of Standard and/or colloquial Russian in that NEOVOC and even more TRADVOC dominate clearly. However, whereas the frequency of the former clearly rises with a higher degree of exposure to Russian, the latter shows decreasing values, though to a limited extent. The Ukrainian vocative variant (TRADVOC) was only marginally represented in the data. The only statement that can be made is that almost all TRADVOC forms stem from respondents with a low exposure to Russian, but even for them TRADVOC is of marginal importance.

The situation with other discourse-organizing expressions, i.e., discourse markers in a broad sense, is different. Here the situation highly resembles the situation of lexical elements in Suržyk in general. In some cases, it is the Ukrainian variant in a pair of functionally equivalent Ukrainian and Russian expressions that may be described as the firmly established default variant in Ukrainian Suržyk, in other cases it is the Russian one, leaving a third group of instances with a considerable degree of variation between the Ukrainian and Russian variant. However, the regional differences do not play such a strong role with discourse markers as, according to Hentschel (2024), is the case with the variation between semantically “full” lexemes of Ukrainian or Russian origin. Among the five hyperlexemes with a considerable variation between Ukrainian and Russian expressions there are two instances of hesitation or disfluency markers which show no correlations with the possible factors tested. For the other three hyperlexemes, the main factor is the respondent’s exposure to Russian: the stronger this is, the higher the probability of using a Russian expression. Only for one of the hyperlexemes (*os’*, *ot*, *osjo* = *vot*) does the regional factor play a role (additionally to exposure to Russian) in that respondents from more Eastern oblasts exhibit a higher probability of using the Russian expression than those from more Western ones.

These instances seemingly reflect a transitional phase before a stabilization in one or the other direction occurs. The question why these five hyperlexemes (still) show variation and the other 11 do not must remain open.

For the stabilized variants of discourse markers of either Ukrainian or Russian origin in Suržyk the question arises of why, for example, it is Ukr. *čuješ* and not Russ. *slyšyš’* ‘[do] you hear’ on the one hand, but on the other hand Russ. *ponimaješ’* and not Ukr. *rozumiješ* ‘[do] you understand’. Explanations for the differing directions of stabilizations like these might be found, if at all, in much deeper analyses of single pairs of competing Ukrainian and Russian discourse markers.

With regard to the borrowability of discourse markers, the data presented here rather confirm doubts concerning whether this internally heterogeneous group of functional expressions can be integrated as a whole into a universal hierarchy. Ukrainian and Russian are genetically structurally close to each other, whereas the vast majority of research on the borrowability of discourse markers deals with borrowing relations between genetically and structurally, if not typologically different languages. As Thomason and Kaufmann (1988) have argued, contact phenomena between genetically and structurally closely-related languages are in many respects different from constellations of the latter type, the former rather obeying the regularities of dialect contact. It is an open question to what extent these regularities in cases like Ukrainian and Russian language contact differ for discourse markers from situations of bilingual contact between two neither genetically closely related nor highly structurally similar languages.

Last but not least, in connection with the internal functional heterogeneity of the class of discourse markers, one further point deserves consideration. Our underlying assumption, formulated at the beginning of this study, was that syntactically isolated expressions are more readily transferred from a socially dominant language (in Ukraine, Russian for more than a century) into the speech of speakers of a socially dominated language. Such syntactically isolated elements might be furthermore more susceptible to succumb to endeavors to imitate a socially dominant language, which Trudgill (1986) distinguishes from accommodation (adaptation) to it.

Taken together, however, our findings do not support these assumptions. Within the class of discourse markers in traditional Suržyk, which is historically based on Ukrainian, some Russian expressions have become established at the expense of their Ukrainian equivalents, whereas in other cases the Ukrainian expressions are equally clearly dominant. This pattern corresponds closely to the findings reported by Hentschel (2024) for nouns, verbs, and adjectives.

In those cases, however, where neither the Ukrainian expression nor its Russian equivalent has yet become established in Suržyk, one observation may be indicative. For two of the hyperlexematic equivalence pairs, *napryklad* = *naprimer* ‘for example’ and *jak by* = *kak by* ‘as if / like’, the Ukrainian expression occurs more frequently in the open interviews than in the family conversations. In contemporary Ukraine, at the latest since the occupation of the Crimean Peninsula by the Russian Federation, Ukrainian has become the unmarked choice

in official and semi-official communicative settings, although, until February 2022 (by which time the data for the present study had already been collected), Russian had by no means become socially discredited (cf. Hentschel & Zeller, 2016). The higher frequency of the Ukrainian expressions from these two equivalence pairs in the open interviews therefore seems to reflect a somewhat stronger orientation toward the language enjoying greater overt prestige in present-day Ukraine.

It would probably be inappropriate to describe this as *imitation*. Rather, it is more plausibly interpreted as a somewhat more conscious choice of linguistic expressions. Notably, however, no comparable pattern can be observed for the hesitation or disfluency markers *ětot* = *cej* and *ěto* = *ce*, derived from forms of corresponding demonstrative pronouns.

After four years of war with the Russian aggressor the “Suržyk-internal” turn to Ukrainian, at least in a more conscious use of languages and codes as in the open interviews postulated above, will most probably be even stronger than in our data collected before 2022. Whether such a turn will affect even those instances of discourse markers or elements in the lexicon in general where the Russian variants are rather firmly established and thus develop Suržyk “backwards” away from Russian to a stronger convergence with Standard Ukrainian, only the future will show. This future will hopefully allow Ukraine to develop into a fully independent and democratic state and an open society that feels itself self-confident enough to allow among many other things a liberal linguistic policy toward minority languages and regional or social subvarieties.

Appendix

(A) On the influence of sociobiographical factors

[5]²⁹ os', ot, osjo = vot

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: *binomial (logit)*

Formula: *Affinity ~ Exposure_to_Russian + Age + Region + (1 | Respondent)*

Categories of *Affinity*: *Russian (Reference)* vs. *Ukrainian*

Number of observations: 5,156³⁰

Random Effects: *Respondent* ($n = 206$, $SD = 1.89$)

Fixed Effects:

	Estimate (β)	SE	z	p
(Intercept)	3.29	1.51	2.17	0.030
Exposure to Russian	-0.14	0.03	-5.24	< 0.001
Age	0.02	0.01	2.58	0.010
Region (Ref.: I)				
II	0.96	1.46	0.66	0.511
III	-1.64	1.42	-1.15	0.251
IV	-0.73	1.42	-0.52	0.606
V	-0.63	1.44	-0.44	0.662

[8] jak by = kak by

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

²⁹ The numbers correspond to those in Table 8.

³⁰ Please note that the numbers of observations (N) reported in the Appendix do not correspond exactly to those given in Table 9 of the main text. The Appendix only includes data from respondents for whom the relevant sociobiographical information is available. Hentschel et al. (forthcoming), in the Introduction to this special issue, explain to what extent—and for what reasons—these data were not collected. Furthermore, the values of N for hyperlexematic pair [5] also differ between Parts (A) and (B) of the Appendix. Respondents for whom sociobiographical data had, in principle, been collected but who did not provide information for a particular parameter were excluded from the analyses involving that parameter.

Family: *binomial (logit)*

Formula: *Affinity ~ Exposure_to_Russian + Subcorpus + (1 | Respondent)*

Categories of *Affinity*: *Russian (Reference)* vs. *Ukrainian*

Number of observations: 895

Random Effects: *Respondent* ($n = 94$, $SD = 6.80$)

Fixed Effects:

	Estimate (β)	SE	z	p
(Intercept)	-0.11	2.37	-0.04	0.965
Exposure to Russian	-0.42	0.14	-2.96	0.003
Corpus: Interview	8.60	2.92	2.95	0.003

[9] napryklad = naprimer

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: *binomial (logit)*

Formula: *Affinity ~ Exposure_to_Russian + Subcorpus + Age + (1 | Respondent)*

Number of observations: 972

Random Effects: *Respondent* ($n = 121$, $SD = 4.02$)

Fixed Effects:

	Estimate (β)	SE	z	p
(Intercept)	-5.39	2.03	-2.66	0.008
Exposure to Russian	-0.18	0.08	-2.17	0.030
Corpus: Interview	3.38	1.27	2.67	0.008
Age	0.07	0.03	2.24	0.025

(B) On interdependencies with inflectional preferences

For all models:

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

Family: *binomial (logit)*

Formula: *Affinity ~ Morphcluster_reduced + (1 | Respondent)*

Categories of *Affinity*: *Russian (Reference)* vs. *Ukrainian*

[5] os', ot, osjo = vot

Number of observations: 4,862

Random Effects: *Respondent* ($n = 162$, $SD = 1.97$)

Fixed Effects:

	Estimate (β)	SE	z	p
(Intercept)	3.09	0.28	11.17	< 0.001
Morphcluster_reduced	-2.35	0.36	-6.48	< 0.001

[6] cej = ètot

Number of observations: 325

Random Effects: *Respondent* ($n = 56$, $SD = 14.99$)

Fixed Effects:

	Estimate (β)	SE	z	p
(Intercept)	10.54	2.13	4.85	< 0.001
Morphcluster_reduced	-2.85	2.26	-1.26	0.209

[7] ce = èto

Number of observations: 377

Random Effects: *Respondent* ($n = 105$, $SD = 9.39$)

Fixed Effects:

	Estimate (β)	SE	z	p
(Intercept)	7.74	1.37	5.64	< 0.001

Morphcluster_reduced	-1.43	1.27	-1.13	0.260
----------------------	-------	------	-------	-------

[8] jak by = kak by

Number of observations: 869

Random Effects: *Respondent* ($n = 86$, $SD = 8.06$)

Fixed Effects:

	Estimate (β)	SE	z	p
(Intercept)	6.63	1.94	3.41	< 0.001
Morphcluster_reduced	-11.54	4.12	-2.80	0.005

[9] napryklad = naprimer

Number of observations: 883

Random Effects: *Respondent* ($n = 100$, $SD = 3.68$)

Fixed Effects:

	Estimate (β)	SE	z	p
(Intercept)	-0.47	0.67	-0.71	0.480
Morphcluster_reduced	-3.56	1.00	-3.57	<0.001

References

- Andersen, G. (2014). Pragmatic borrowing. *Journal of Pragmatics*, 67, 17–33.
- Auer, P. (1999). From code switching via language mixing to fused lects: Towards a dynamic typology of bilingual speech. *International Journal of Bilingualism*, 3(4), 309–332.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bevzenko, S. P. (1980). *Ukraïns'ka dialektolohija*. Vyšča škola.
- Bevzenko, S. P., Hryščenko, A. P., & Lukinova, T. B. (Eds.). (1978). *Istoriya ukraïns'koï movy. Morfolohija*. Naukova dumka.
- Blankenhorn, R. (2003). *Pragmatische Spezifika der Kommunikation von Russlanddeutschen in Sibirien. Entlehnung von Diskursmarkern und Modifikatoren sowie Code-switching*. Peter Lang.
- Bogdanova-Beglarian, N., Blinova, O., Sherstinova, T., Martynenko, G., & Zajdes, K. (2018). Pragmatic Markers in Russian Spoken Speech. In S. Balandin, T. S. Cinotti, F. Viola, & T. Tyutina (Eds.), *Proceedings of the 23rd Conference of Open Innovations Association (FRUCT)* (pp. 17–24). FRUCT Oy. <https://doi.org/10.23919/FRUCT.2018.8588101>
- Burgsteden, L. van (2023). Generalized list completer. In A. Gubina, E. M. Hoey, & C. W. Raymond (Eds.), *Encyclopedia of Terminology for Conversation Analysis and Interactional Linguistics*. International Society for Conversation Analysis (ISCA). <https://doi.org/10.17605/OSF.IO/4URZQ>
- Chambers, J. K. (2002). Patterns of variation including change. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 349–372). Malden Blackwell.
- Crible, L. (2018). *Discourse markers and (dis)fluency: forms and functions across languages and registers*. (= *Pragmatics and Beyond*. New Series, 286). John Benjamins Publishing Company.
- Daniel', M. A. (2009). "Novyj" russkij vokativ: istorija formy usečennogo obraščenija skvoz' prizmu korpusa pis'mennyx tekstov. In K. L. Kiseleva, V. A. Plungyan, E. V. Rahilina, & S. G. Tatevosov (Eds.), *Korpusnye issledovanija po russkoj grammatike*. Probel.
- Del Gaudio, S. (2010). *On the nature of Suržyk: a double perspective*. [Wiener slawistischer Almanach, Sonderband 75]. Verlag Otto Sagner.

- Dingemanse, M. (2024). Interjections at the Heart of Language. *Annual Review of Linguistics*, 10, 257–277. <https://doi.org/10.1146/annurev-linguistics-031422-124743>
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer.
- Gorbačevič, K. S. (Ed.). (2004 –). *Bolšoj akademičeskij slovar' russkogo jazyka*. Nauka.
- Gradus (2026). *Doslidžennja mediaspožyvannja naseleennja Ukraïny: Analitičnyj zvit (ljutyj–berezennja 2026)*. Gradus Research Plus. https://mincult.gov.ua/wp-content/uploads/2026/04/gradus_report_final-present_mincult_2004206.pdf
- GRAK: Švedova, M., fon Val'denfel's, R., Jaryhin, S., Rysin, A., Starko, V., Nikolajenko, & T., Lukaševs'kyj, A. (2017–). *Heneral'nyj rehional'no anotovanyj korpus ukraïns'koï movy*. <https://uacorporus.org/>
- Haugen, E. (1950). The analysis of linguistic borrowing. *Linguistics*, 26, 210–231.
- 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 XVI. Internationalen Slavistenkongress Belgrad 2018* (Reihe Die Welt der Slaven. Sammelbände—Sborniki) (pp. 127–138). Harrassowitz Verlag.
- Hentschel, G. (2024). Ukrainian and Russian in the lexicon of Ukrainian Suržyk: Reduced variation and stabilisation in Central Ukraine and on the Black Sea coast. *Russian Linguistics*, 48(2), article 2. <https://doi.org/10.1007/s11185-023-09286-9>
- Hentschel, G. (forthcoming a). Hybridization from two sides? Ukrainian- vs. Russian-based Suržyk. Historical and theoretical aspects and empirical insights. *Russian Linguistics*, 50.
- Hentschel, G. (forthcoming b). Inflectional morphology in Ukrainian-Russian code-mixing—On endings and stems of main lexical word classes in Suržyk. *Russian Linguistics*, 50.
- Hentschel, G., & Palinska, O. (2022). Restructuring in a mesolect: A case study on the basis of the formal variation of the infinitive in Ukrainian-Russian “Suržyk”. *Cognitive Studies | Études cognitives*, 22, Article 2770. <https://doi.org/10.11649/cs.2770>
- Hentschel, G., & Taranenko, O. (2021). Bilingualism or tricolalism: Ukrainian, Russian and “Suržyk” in Ukraine. Analysis and linguistic-geographical mapping. *Die Welt der Slaven*, 66(2), 268–299. <https://doi.org/10.13173/WS.66.2.268>
- Hentschel, G., & Zeller, J. P. (2016). Meinungen und Einstellungen zu Sprachen und Kodes in zentralen Regionen der Ukraine. *Zeitschrift für Slawistik*, 61(4), 636–661.
- Hentschel, G., Palinska, O., & Zeller, J. P. (forthcoming). Introduction: background, aims, source material, and brief outlook on this special issue on Ukrainian Suržyk. [Special issue “Suržyk” between Ukrainian and Russian – Hybridization from Two Sides?]. *Russian Linguistics*, 50.
- Il'in, V. S., & Dorošenko, S. I. (1965). *Ukraïns'ko-rosijs'kyj slovnyk / Ukraïnsko-russkij slovar'*. Naukova dumka.
- Isačenko, A. V. (1975 / 1982). *Die russische Sprache der Gegenwart. Formenlehre*. Hueber.
- Źakevyč, H., Kaljužna, V., Palamarčuk, O., Pylins'kyj, M., Skrypnyk, & L., Čertoryz'ka, T. (2004). *Ukraïns'ko-rosijs'kyj slovnyk / Ukraïnsko-russkij slovar'*. Naukova dumka.
- Jakobson, R. (1962). Why ‘mama’ and ‘papa’? In R. Jakobson, *Selected Writings* (Vol. I: Phonological Studies, pp. 538–545). Mouton.
- Janda, L. A. (2019). Name-calling: The Russian ‘new Vocative’ and its status. In J. C. Salmons (Ed.), *Current issues in linguistic theory* (pp. 381–394). John Benjamins. <https://doi.org/10.1075/cilt.345.18jan>
- Kobčenko, N. (2021). Vokatyv i nominativ u funkciji zvertannja: formuvannja movnoji normy v radjans'kyj period. *Mova: klasyčne—moderne—postmoderne*, 7, 68–91. <https://doi.org/10.18523/lcmp2522-9281.2021.7.68-91>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(1), 1–26. <https://doi.org/10.18637/jss.v082.i13>

- Lukashevskiy, A., Zakharov, K., & Shvedova, M. (2025). *Rada_Trees: A Syntactically Annotated Corpus of Ukrainian Parliament Transcripts (1990-2024)*. https://huggingface.co/datasets/uacorpora/Rada_Trees
- Matras, Y. (2007). The borrowability of structural categories. In Y. Matras, J. Sakel (Eds), *Grammatical Borrowing in Cross-Linguistic Perspective* (pp. 31–73). De Gruyter Mouton. <https://doi.org/10.1515/9783110199192.31>
- Matras, Y. (2009). *Language contact*. Cambridge University Press.
- Matvijas, I. H. (1974). *Imennyk v ukra'ns'kij movi*. Radjans'ka škola.
- Mazničenko, J., Makedon, V., Šarabanova, S., & Jalovnyča, I. (Eds.). (2019). *Ukraïns'kyj pravopys*. Naukova dumka.
- Mišić Ilić, B. (2017). Pragmatic borrowing from English into Serbian: Linguistic and sociocultural aspects. *Journal of Pragmatics*, 113, 103–115.
- Muysken, P. (2000). *Bilingual speech. A typology of code-mixing*. Cambridge University Press.
- NKRJa: *Nacional'nyj korpus russkogo jazyka*. <https://ruscorpora.ru> (cf. Savčuk et al., 2024)
- Norrick, N. R. (2007). Discussion article: Pragmatic markers, interjections and discourse. *Catalan Journal of Linguistics*, 6(1), 159–168. <https://doi.org/10.5565/rev/catjl.129>
- Norrick, N. R. (2009). Interjections as pragmatic markers. *Journal of Pragmatics*, 41, 866–891. <https://doi.org/10.1016/J.PRAGMA.2008.08.005>
- OCUS-Doc: Hentschel, G., Zeller, J. P., & Palinska, O. (2026). *Documentation of the “Oldenburg Corpus of Ukrainian Suržyk”*. <https://uol.de/slavistik/forschung/sprachwissenschaft/bereich-mischvarieteten/okus-dok>
- Palinska, O., & Hentschel, G. (2022). Regional'nye osobennosti ispol'zovanija ukrainsko-russkoj smešannoj reči (suržika) i vlijanie dialektov: pristavki i predlogi VID / OT. *LingVaria*, 2(34), 229–253. <https://doi.org/10.12797/LV.17.2022.34.15>
- Peterson, E. (2017). The nativization of pragmatic borrowings in remote language contact situations. *Journal of Pragmatics*, 113, 116–126.
- Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics*, 18(7–8), 581–618. <https://doi.org/10.1515/ling.1980.18.7-8.581>
- Proske, N. (2014). °h ach KOMM; hör AUF mit dem klEInkram. Die Partikel *komm* zwischen Interjektion und Diskursmarker. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 15, 121–160.
- R Core Team (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Savčuk, S. O., Arxangel'skij, T. A., Bonč-Osmolovskaja, A. A., Donina, O. V., Kuznecova, Ju. N., Ljaševskaja, O. N., Orexov, B. V., & Podrjadčikova, M. V. (2024). Nacional'nyj korpus russkogo jazyka 2.0: novye vozmožnosti i perspektivy razvitija. *Voprosy jazykoznanija*, 2(7), 7–34. <https://ruscorpora.ru/page/corpora-quote>
- Šerex, Ju. (1951). *Narys sučasnoï ukrains'koï literaturnoï movy*. Molode žyttja.
- Shvedova, M., & Lukashevskiy, A. (2025). Case choice in Ukrainian vocative expressions: A study of parliamentary transcripts (1990–2024) annotated with universal dependencies. In A. Kalnača, P. Ozola, & M. Hoplíček (Eds), *Grammar and corpora. 10th international conference. Book of abstracts* (pp. 106–108). University of Latvia Press.
- Sonnenhauser, B., & Hanna, P. N. A. (2013). Introduction: Vocative! In B. Sonnenhauser, & P. N. A. Hanna (Eds.). *Vocative! Addressing between System and Performance* (pp. 1–23). De Gruyter.
- Stifter, D. (2013). Vocative for nominative. In B. Sonnenhauser, & P. N. A. Hanna (Eds.). *Vocative! Addressing between System and Performance* (pp. 43–86). De Gruyter.

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). Peter Lang.

Taranenko, O. (2007). Ukrainian and Russian in contact: attraction and estrangement. *International Journal of the Sociology of Language*, 183, 119–140.

Taranenko, O. (2024). *Ukrains'ka literaturna mova kincja XX—peršoji čverti XXI st.: stan i tendencii rozvytku*. Vol. II. Ukrains'kyj movno-informacijnyj fond NAN Ukraïny.

Theodoropoulou, I. (2009). Discourse markers vs. interjections: An interactional sociolinguistic approach. *Language Forum: A Journal of Language and Literature*, 35(1), 79–97.

Thomason, S. G. & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.

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

Unuabonah, F. O. & Mabena, M. I. (2024). Discourse-Pragmatic Borrowing in South African English. *Language Matters*, 55(1–2), 70–91.

Žajvoronok, V. V. (Ed.). (2003). *Rosijs'ko-ukraïns'kyj slovnyk. A-Ja / Russko-ukrainskij slovar'. A-Ja*. Abrys.

Zaliznjak, A. (1973). O ponimannii termina “padež” v lingvističeskix opisanijax. In A. Zaliznjak (Ed.), *Problemy grammatičeskogo modelirovanija* (pp. 53–87). Nauka.

Zeller, J. P. (forthcoming). The impact of Ukrainian and Russian on sound patterns in Suržyk. *Russian Linguistics*, 50.

Zwicky, A. (1974). “Hey, Whatsyourname”. In M. La Galy, R. A. Fox, & A. Bruck (Eds.), *Papers from the Tenth Regional Meeting, Chicago Linguistic Society. April 19–21, 1974* (pp. 787–801). Chicago Linguistic Society.