OV/VO Word Order in Heritage Russian: Is Transfer at Play?

Maria Martynova, Yulia Zuban, Luka Szucsich, and Natalia Gagarina

Abstract: The present study investigates the choice of OV/VO word order in heritage and monolingual Russian. In monolingual Russian, OV/VO order is claimed to be sensitive to the object realization (noun vs. pronoun) and clause type (main vs. embedded). In heritage Russian, OV/VO order is claimed to be prone to changes under language contact. Analyzing spoken and written narratives produced by heritage speakers (HSs) of Russian residing in the US and Germany, we scrutinize HSs' choice of OV/VO orders in comparison to the monolingual speakers from Russia. According to the results of the binomial generalized linear mixed-effects model, the OV/VO choice in heritage Russian was best predicted by the clause type and object realization. Specifically, the likelihood of producing the OV order was lower in the embedded clauses than in the main clauses among all speaker groups. Furthermore, all three speaker groups preferred the OV order with the pronominal object, while the preference shifted towards the VO order when the object was realized by a noun. Finally, both HS groups behaved similarly to the monolingual speakers in their choice of OV/VO orders. The results of the study do not provide any clear evidence for cross-linguistic influence from the majority languages and suggest that the word order choice of heritage and monolingual speakers depends on multiple factors, such as clause type and object realization.

1. Introduction*

The number of migrants worldwide increases continuously, establishing robust language contact situations in receiving countries. Some countries, such as the US and Germany, have a substantial proportion of immigrants, around 15.5% and 16% of the population, respectively (United Nations 2019).

* This research was funded by the German Research Foundation (DFG) as part of the research unit "Emerging grammars in language contact situations: A comparative approach" (FOR 2537), project no. 313607803, GZ: ZE 940/4-1, SZ 263/6-1, GA 1424/10-1.

In the US, approximately 880,000 speakers report that they speak Russian at home, making it the ninth most spoken family language (US Census Bureau 2015). Russian-speaking communities in the US are typically not located in particular neighborhoods but rather are scattered across the country (Dubinina and Polinsky 2013; Laleko 2013). Nevertheless, there are some places with a higher density of Russian speakers. The largest Russian-speaking communities are situated in New York (226,290 speakers), California (151,685), and Washington (56,537) (US Census Bureau 2015).

In Germany, approximately 3.487 million people originate from the countries of the former Soviet Union (Statistisches Bundesamt 2019a, 2019b). However, the real number of Russian speakers in Germany may reach up to 6 million people (cf. Anstatt 2011; Aref'ev 2012; Brehmer 2007; Brehmer and Mehlhorn 2015; Gagarina 2014, 2017; Gagarina and Klassert 2018; Hamann et al. 2019; Topaj 2018).

The present study focuses on heritage Russian that is spoken in the US and Germany. We adopt the following definition of HSs and heritage languages (HLs): "Heritage languages are languages spoken by the children of immigrants or by those who immigrated to a country when young" (Cho et al. 2004: 23). HSs are therefore bi- or multilingual speakers, using their HL(s) alongside the majority language(s) (ML) of the surrounding community (cf. among many others, Benmamoun et al. 2013; Guijarro-Fuentes and Schmitz 2015; Montrul 2015; Polinsky 2015; Rothman 2009; Valdés 2005).

The language performance of HSs of Russian was often reported to differ from the monolingual speakers of Russian (Laleko 2019; Romanova 2008). Compared to the monolingual speakers, heritage Russian was found to show salient changes in lexicon (Isurin 2011; Polinsky 2006), nominal morphosyntax such as case and gender agreement (Laleko 2018; Polinsky 2006, 2008a), verbal morphology (Romanova 2008), verbal aspect (Gagarina et al. 2020; Laleko 2010, 2011, 2015; Polinsky 2006, 2008b), pro-drop (Dubinina and Polinsky 2013; Isurin 2011; Laleko and Polinsky 2017), and word order (Brehmer and Usanova 2015; Dubinina and Polinsky 2013; Isurin and Ivanova-Sullivan 2008; Kisselev 2019; Laleko and Dubinina 2018; Polinsky 2006, 2011; Zuban et al. 2021).

Although word order in heritage Russian has been investigated in a number of studies, none of them, to our knowledge, explicitly focused on the OV/VO orders in the narrow sense, i.e., the ordering of objects/internal arguments relative to the main verb. The present study aims to fill in this gap. Grouping the data into OV/VO orders allows us to include the dataset by setting minimal exclusion criteria, as specified in §4.3. Limiting ourselves to the OV/VO distinction is also grounded in its importance for the theoretical discussion on head directionality in syntactic structures in general (cf. for a discussion, among many others, Baker 2001; Broekhuis 2006; Haider 2015; Haider and Szucsich 2022; Kayne 1994).

OV/VO orders represent a promising field for investigation in language contact for a number of reasons. The linearization of objects (internal arguments) relative to the main verb—especially in languages with so-called free word order—is determined not only by the directionality (licensing of verbal complements by initial or final heads), but also by information-structural properties of the utterance (see Bailyn 1995; Junghanns and Zybatow 1995; Kallestinova 2007; Slioussar 2011, among many others for Russian). Head-directionality is associated with core syntax (Aboh 2015; Uriagereka 2007)[1] and is considered to be less susceptible to contact-induced and diachronic change (Kroch and Taylor 2000; Pintzuk and Taylor 2006). Linearization of constituents due to information structure, on the other hand, concerns interfaces of core grammar with discourse. Furthermore, in monolingual Russian, OV/VO choice is highly influenced by object realization (noun vs. pronoun) (Kallestinova 2007).

In light of the Interface Hypothesis (see Sorace 2011 for a detailed discussion), one can come up with two competing predictions for languages with free word order: (i) If one assumes that head-directionality is the determining factor for the distribution of OV/VO orders, one would expect it to be less prone to contact-induced change resulting in a higher degree of optionality; and (ii) if one assumes that information structure is the decisive factor, one would expect the distribution of OV and VO in HSs to deviate from monolingual production by showing a higher degree of variability under constant conditions (object type: nominal vs. pronominal, and clause type: main vs. embedded).

Since the Interface Hypothesis is difficult to test, the aim of the current study is more modest, viz., to investigate the possible influence of the MLs, English and German, on the choice of OV/VO orders in heritage Russian by comparing two HS groups to monolingual speakers of Russian. The data of the participants are analyzed with the generalized mixed-effects modeling that accounts for the individual speaker variation and allows inclusion of other factors that influence the choice of OV/VO orders (they will be specified in detail later).

The paper is organized as follows. Section 2 provides theoretical background on verb placement in monolingual English, German, and Russian. Section 3 specifies the research questions alongside the hypotheses, derived from the previous studies on word order in heritage and monolingual Russian. Section 4 describes the experimental design, participants, and data analysis. Section 5 presents the results of the study, which are discussed in §6.

[1] Of course, this is true only if parametrized directionality is accepted within syntactic theory; for an alternative, see Kayne's (1994) proposal of the Linear Correspondence Axiom, which universally allows only for head-initial derivations.

2. Theoretical Background

2.1. Verb Placement in Majority Languages English and German

English is an SVO language with residual V2, e.g., quotative inversion, and very limited reordering of constituents, e.g., topicalization of objects (Birner and Ward 1998; Eppler 1999; Kempen and Harbusch 2019). This holds irrespective of what we call "clause type" in this article, i.e., it is valid for both main and embedded clauses, as illustrated in the example below:

(1) My friends know that I like this book.

In contrast to English, German is an SOV language with V2 in main clauses and reordering options for non-verbal constituents, i.e., German is well-known for exhibiting scrambling (Eppler 1999; Gärtner 2000; Haider 2012; Kempen and Harbusch 2019; Wegener 1993). In embedded clauses, introduced by complementizers and relativizers/wh-elements, verbal categories marked for finiteness—finite main verb or finite auxiliaries—canonically occupy the clause-final position (with the exception of particular, so-called non-integrated embedded clauses introduced by *weil* 'because', *obwohl* 'although', etc.; see Gärtner 2000 and Wegener 1993):

(2) Ich mag das Buch, das ich zum Geburtstag bekam.
I like the book which I for birthday got
'I like the book that I got for my birthday.'

Thus, German makes a clear distinction between main and embedded clauses when it comes to surface order, with robust OV in embedded and predominant VO order in main clauses with synthetic verb forms and OV with analytic verb forms. In the generative tradition, the V2 effect is derived by movement of the finite verbal element to the C-head, i.e., it is a grammaticalized process (Gärtner 2000 and Haider 2012, among many others).

2.2. Verb Placement in Monolingual and Heritage Russian

Monolingual Russian is considered to have an SVO order with full (non-pronominal) object NPs in neutral contexts, i.e., with sentential ("all new") and VP focus in both main and embedded clauses. At the same time, Russian is a language which allows massive re-orderings/permutations of constituents (Brehmer and Usanova 2015; Junghanns and Zybatow 1995; Kallestinova 2007; Slioussar 2007, 2011; Švedova 1980, 2005). In the generative literature on Russian, it is widely assumed that objects are "base-generated" to the

right of the verbal head (Bailyn 1995; Jasinskaja and Šimík, forthcoming; Junghanns and Zybatow 1995; Slioussar 2007, 2011).[2] One of the arguments for this assumption is that the VO order is by far the most frequent relative order in clauses with a verb selecting an internal argument (object). Bailyn (2012: 7) states that Russian has a head-initial VP. The vast majority of authors assume that the above-mentioned positional alternations are governed by information structure, but crucially they are also taken to be derived from the basic/neutral structure by additional operations (movement), although the exact motivation and landing sites of these movements might differ (Bailyn 1995, 2004; Junghanns and Zybatow 1995; and Slioussar 2007, among others). The following example contains both basic and derived word orders:

(3) Context: [The very beginning of a text] On November 29, 2018, an accident occurred at the address…

Mužčina	uronil	mjač	na	dorogu	za	nim	pobežala	sobaka.
man	dropped	ball	on	road	after	him	ran	dog

'A man dropped a ball on the road and a dog ran after it.'

(RUmo65FR_fwR,[3] RUEG Corpus)

In example (3) from a narration of a monolingual speaker, the SVO word order is used in an "all-new" context at the beginning of the story, while the following OVS order is used to introduce a new subject, a dog, into the discourse (see, e.g., Junghanns and Zybatow 1995; Kallestinova 2007; and Slioussar 2007 for detailed discussion). Since we will investigate VO vs. OV ordering in Russian quantitatively, a qualitative assessment of the data from an information-structural perspective must be left for further research. Although word order in monolingual Russian (under appropriate information structural conditions) is claimed to vary independently of the clause type, the variation is expected to be higher in the main clause than in the embedded one (Bailyn 2012). Zuban et al. (2021) showed that monolingual speakers of Russian produced fewer word order combinations in embedded clauses (namely, five) compared to the main clauses (namely, six). Moreover, the frequency of the SVO order produced by monolingual speakers was higher in embedded clauses (around 62%) compared to the main clauses (roughly 56%) (Zuban et al. 2021).

In addition to the information structure, word order in monolingual Russian is influenced by the object realization. Specifically, there is a strong

[2] But for an alternative view, see Haider and Szucsich (2022), in which the authors argue that Slavic languages are ambivalent between SVO and SOV (so-called T3-languages).

[3] The meaning of the participants' code is provided in Table 3.

tendency for the pronominal objects to be placed preverbally (Kallestinova 2007), as shown in example (4):

(4) Context: The first driver braked and that was it.

Vtoroj v nego vrezalsja.
second in him bumped

'The second one bumped into him.' (RUmo24FR_iwR)

Additionally, word order in monolingual Russian is influenced by other factors that will not be investigated in our study (e.g., intonation, grammatical weight of constituents, etc.; for more detailed information, see Bailyn 2012; King 1995; Laleko and Dubinina 2018; Lobanova 2011; Luchkina and Cole 2016).

Word order in HLs is claimed to be prone to change possibly induced by transfer. As a result, in languages with multiple word order options, the word order variation is reduced (Polinsky 2018). Such reduction of word order options is shown in several studies on different HLs such as heritage Korean (O'Grady et al. 2011), heritage Norwegian (Johannessen and Laake 2015), heritage Spanish (Cuza 2012; Cuza and Frank 2011), and heritage Hungarian (Fenyvesi 2005). However, Polinsky (2018: 273) points out that the word order reduction might not be a general outcome of language contact. Since the majority of studies focused on the populations whose ML is English, the behavior of HSs with a different ML is less predictable.

Several studies on word order in heritage Russian in the US found that HSs reduced their word order flexibility and increased the use of the SVO orders (e.g., Isurin 2005; Kisselev 2019; Laleko and Dubinina 2018; Polinsky 2006; Zuban et al. 2021).

Polinsky (2006: 237) states that HSs of Russian in the US clearly limit their word order repertoire by mainly producing SVO order with both nominal and pronominal arguments. Although Polinsky (2006) reports on the "frozen" word orders of HSs of Russian in the US, it has to be noted that these observations are not based on any numerical analysis of the word order patterns.

Isurin (2005) conducted a longitudinal study with one Russian girl who came to the US at the age of nine. The girl took part in different tasks that were repeated after a 9-month break. The results revealed that during the second elicitation session the girl produced more SVO and fewer inverted word orders compared to the first session. The results of the study were explained with reference to transfer from the ML and language internal factors such as frequency of inverted and basic word orders in the HL (for more details, see Isurin 2005: 1122).

Laleko and Dubinina (2018) investigated word order patterns in the "frog stories" of 21 HSs of Russian in the US and 19 monolingual speakers of Russian. The results of the study showed that HSs produced significantly more SVO

orders and significantly fewer non-SVO orders than the monolingual speakers (Laleko and Dubinina 2018: 202). Laleko and Dubinina (2018: 202) conclude that the relative flexibility of word order typical for monolingual Russian is reduced in HSs of Russian in the US.

Another study by Kisselev (2019) examined word order patterns of HSs of Russian in the US in written essays. It was found that HSs produced more SV and fewer VS orders than monolingual speakers. However, SV orders not only included the SVO order, but also all other word orders in which the verb follows the subject (i.e., SOV and OSV). Furthermore, although HSs were found to produce various word order combinations, they did not produce any verb-initial word orders, contrary to the monolingual speakers who produced all six grammatically possible word order combinations (Kisselev 2019: 161).

A more recent study by Zuban et al. (2021) examined the word order patterns of three speaker groups, namely HSs of Russian residing in the US and Germany, and monolingual speakers of Russian residing in Russia. Word order was investigated in a corpus of semi-spontaneous spoken and written narratives collected according to the "Language Situations" method (Wiese 2020). First, only the results related to the HSs in the US and monolingual speakers will be summarized, while the HSs in Germany will be discussed later. The results of the study showed that HSs residing in the US were found to produce significantly more SVO and significantly fewer OVS orders than monolingual speakers in the overall dataset. Furthermore, the split of the data into main and embedded clauses revealed further interesting findings. In the main clauses, both the HSs and monolingual speakers behaved similarly regarding the overall distribution of different word orders. In the embedded clauses, however, the HSs significantly differed from the monolingual speakers by predominantly producing the SVO order. Although monolingual speakers were found to increase the rate of SVO order in embedded clauses compared to main ones, that increase was not as substantial (from around 56% to 62%) as in the case of the HSs (from around 65% to 91%). Also, the HSs in the US produced fewer word order combinations in the embedded clauses (four combinations) than in the main ones (six combinations) (Zuban et al. 2021: 270).

Although word order in heritage Russian was often investigated for HSs with the ML English (e.g., Kisselev 2019; Laleko and Dubinina 2018; Polinsky 2006; Zuban et al. 2021), much less is known about HSs with the ML German. In one of a few studies on heritage Russian word order in bilingual speakers in Germany, Brehmer and Usanova (2015) examined word order patterns in written productions of HSs of Russian residing in Germany. Their study aimed at investigating the potential influence of German on the verb placement in declarative sentences in heritage Russian. The data were elicited from 20 HSs of Russian in Germany and 20 monolingual speakers from Russia. All participants were asked to write two texts representing

two different situational settings: a magazine article about how to create a boomerang (task 1: "Fast Catch Boomerang" (more formal academic language was expected)) and a story about how a boomerang can be used (task 2: "Boomerang in the Park" (elements of narrative speech were expected)).

The results of the study showed that HSs did not significantly differ from the monolinguals regarding the placement of V2 in main clauses (Brehmer and Usanova 2015: 174). As for the V-final position in embedded clauses, HSs produced significantly more V-final linearizations than monolingual speakers in narrative texts, but more formal texts did not differ across the two speaker groups (Brehmer and Usanova 2015: 179). Besides, in main clauses HSs also produced V-final orders significantly more frequently than monolingual speakers (for main clauses, Brehmer and Usanova (2015: 180) did not distinguish between different tasks in their analysis of main clauses). The results of the V-final position in embedded clauses were interpreted as language transfer from German to Russian, while the results of the V-final placement in main clauses were explained with reference to pragmatic unmarking (for details on pragmatic unmarking, see Brehmer and Usanova 2015: 182). Since Brehmer and Usanova (2015) do not explicitly focus on OV/VO orders, one can only derive the frequency of OV/VO from all word order patterns (OV: OV, SOV, OSV, OVS vs. VO: VO, SVO, VOS, VSO) (Brehmer and Usanova 2015: 179). However, Brehmer and Usanova (2015) do not report the frequencies of different word order patterns in embedded clauses. Thus, the distribution of OV and VO given in Table 1 is restricted to main clauses. It can be seen that HSs produced OV orders more frequently than monolingual speakers.

Table 1. Aggregated absolute and relative distributions of OV/VO patterns in Brehmer and Usanova (2015) in main clauses

Group	OV vs. VO, task 1	OV vs. VO, task 2
Bilinguals	31 (28.2%) vs. 79 (71.8%)	9 (18.4%) vs. 40 (81.6%)
Monolinguals	7 (4.6%) vs. 146 (95.4%)	2 (4.1%) vs. 47 (95.9%)

The study by Zuban et al. (2021), which was mentioned above, also investigated the word order patterns produced by HSs of Russian in Germany. The results of the study, contrary to the results of Brehmer and Usanova (2015), revealed that HSs in Germany showed no significant difference from monolingual speakers regarding the word order patterns produced in both main and embedded clauses. Furthermore, the study showed that clause type had an effect on the word order repertoire and on the frequency of SVO word order. Specifically, both heritage and monolingual speakers produced six word or-

der combinations in the main clauses and five word order combinations in the embedded clauses. Besides that, both speaker groups increased the frequency of SVO order in the embedded clauses compared to the main ones.

To sum up, the results regarding word order in heritage Russian in Germany are conflicting. On one hand, HSs were found to differ from monolingual speakers by producing more V-final linearizations in both main and embedded clauses (Brehmer and Usanova 2015). On the other hand, HSs were found to pattern with the monolingual speakers regarding the production of different word order patterns (Zuban et al. 2021). The studies on word order in heritage Russian in the US typically report that the HSs increase the proportion of SVO orders and decrease word order variation (Kisselev 2019; Laleko and Dubinina 2018). Interestingly, clause type was found to be a major factor that can influence the word order choice of the HSs (Zuban et al. 2021). However, the aforedescribed studies differ from each other in terms of methodology (e.g., tasks, elicitation techniques, participant sample, data annotation, and analysis), and this factor might have contributed to the divergent results of the studies.

3. Research Questions and Predictions

Taking the previous studies on word order in heritage Russian into consideration, we aim at contributing to our understanding of whether OV and VO orders are subject to transfer from the ML of English or German. The current study seeks to answer the following research questions:

RQ1 Do we find a crosslinguistic influence from the ML English on the choice of OV/VO orders in heritage Russian?

H1 HSs are expected to produce VO orders more frequently than monolingual speakers regardless of the object realization in both main and embedded clauses (Isurin 2005; Kisselev 2019; Laleko and Dubinina 2018; Polinsky 2006; Zuban et al. 2021).

RQ2 Do we find a crosslinguistic influence from the ML German on the choice of OV/VO orders in heritage Russian?

H2 HSs are expected to produce OV orders more frequently than monolingual speakers in embedded clauses with nominal objects (Brehmer and Usanova 2015).

4. Methodology

The present study analyzes 96 narrations drawn from the RUEG corpus (Wiese et al. 2021). The RUEG corpus was created within the Research Unit "Emerging Grammars in Language Contact Situations: A Comparative Approach"

(RUEG). RUEG investigates the linguistic systems and linguistic resources of HSs of different MLs and HLs (German, English, Greek, Russian, and Turkish). This article focuses on the narrations of the HSs of Russian. In §4.1 and §4.2, we will describe in detail the design of the RUEG subcorpus for Russian (RUEG-RU), including the data elicitation method, an overview of parameters and factors controlled, participants' coding and metadata, and available annotation.

4.1. Experimental Design

The data were elicited according to the Language Situations Method (Wiese 2020). The participants were shown a video of a fictional car accident. Multiple characters were involved in the accident: a young woman with a stroller accompanied by her son with a ball, another woman with a small dog who was putting her groceries in the trunk of her car, and two drivers (of a white car and of a blue car). The action unfolds as follows: the boy suddenly drops the ball, the dog runs after it, the cars stop abruptly, and the second car slightly bumps into the first one. Additionally, the woman drops a paper bag with her groceries, and they roll away. At the end of the story, everybody is safe, the boy helps the woman to collect her groceries, and the two car drivers call the police. After the participants watched the video, they were asked to narrate what they saw in four different communicative situations that differed with respect to their formality and mode:

1. **Informal spoken**: a voice message to a friend via WhatsApp
2. **Informal written**: a text message to a friend via WhatsApp
3. **Formal spoken**: a witness call to the police
4. **Formal written**: a written witness report to the police

HSs had two elicitation sessions, one in their HL and one in their ML. These sessions were separated from each other by at least three days. Each elicitation was performed by two different elicitors, one for the formal and one for the informal situation. The elicitations took place in two different rooms: one room looked more formal (similar to a typical office or a school room), while another room looked more informal (with some tablecloths, snacks, and juice). The participants had a break between the formal and the informal elicitations. Besides this, the order of four different narrations was randomized. Finally, the participants filled out a background questionnaire about their language use and input, family, and socio-economic status.

4.2. Participants

Table 2 gives detailed information about the participants that formed part of the current study: country of elicitation, number of participants, speakers' gender, number of narrations, speakers' age per group, and age of onset (AoO) of the ML.[4]

Table 2. Participants

Country	Speakers	Females	Narrations	Mean age	AoO ML
Germany	8	4	32	17.0, SD = 0.84	3.1, SD = 1.3
Russia	8	4	32	16.6, SD = 0.49	n/a
US	8	4	32	15.7, SD = 1.51	4.2, SD = 2.26

The participants in the US were recruited in the greater Washington area mainly with the help of the Metaphor Academic Center for Russian Language & Culture in Virginia. The participants in Germany were recruited in the Berlin and Brandenburg urban area via calls in mailing lists, social media, educational institutions (e.g., schools, universities, and language courses), and public organizations (libraries, youth and sport clubs, shopping centers, and consulting offices). The HSs of Russian in the US and Germany were invited to participate if they grew up in one of these two countries speaking Russian with at least one Russian-speaking caregiver at home. Only candidates with one HL were admitted to the experiment.

All HSs were either born in the US or Germany or moved there before the age of six. Furthermore, all HSs in our sample still attended school or just finished school, but more importantly, still lived with their Russian-speaking caregivers at the moment of testing. In addition, none of the HSs attended bilingual schools, but some of the participants learned Russian in a language school (one HS in the US) or in school (five HSs in Germany) as a foreign language. Almost all participants could write in the Russian Cyrillic

[4] The participants' sample used in the present study is the same as the one in the study by Zuban et al. (2021). However, in contrast to Zuban et al. (2021), the present study investigates another syntactic phenomenon involving a different set of data. Specifically, Zuban et al. (2021) investigated six trivalent word order patterns (SVO, SOV, VOS, VSO, OSV, OVS), whereas the present study analyzes OV and VO word orders irrespective of the subject position if applicable. Also, a more advanced statistical modeling allowing us to control for several factors and a different annotation scheme were applied in the present study.

script.[5] Also, four HSs in the US and two HSs in Germany participated in other Russian-speaking activities such as music lessons, dance classes, etc. As can be seen in Table 2, the participants in both HSs groups are quite homogeneous regarding the AoO in their ML, which is also true for their language background in Russian, although HSs in Germany may have slightly more exposure to formal education in Russian than those in the US.

As for the ML of the HSs, some of the participants (a total of three HSs in the US and Germany) started acquiring their ML from birth, while other HSs started acquiring their ML later than their HL. As can be seen in Table 2, on average HSs in the US started to acquire their ML slightly later than HSs in Germany, namely, at the age of 4.2 and at the age of 3.1, respectively.

In the background questionnaire, the bilingual participants were asked about their use of different types of media in their HL and ML, including video, audio, writing, and reading. The HSs of both groups reported usage of the media more frequently in their ML than in their HL, suggesting that majority English and German are the main languages in their daily life.

Next to the self-reports about the use of different media, the HSs assessed their knowledge of Russian in four language domains: understanding, speaking, reading, and writing. The HSs of both groups reported that it was fairly easy for them to understand, speak, and read in Russian. As for the writing, while the HSs in the US found it easy to write in Russian, the HSs in Germany found it more difficult.

Finally, the HSs were asked about the use of their HL and ML with their core family, i.e., which language/languages is/are typically used to communicate with their parents. According to their self-reports, the HSs in the US were exposed to Russian at home to a lesser degree than the HSs in Germany.

Monolingual participants were recruited in St. Petersburg, Russia, primarily through calls in social media and educational institutions. Monolingual speakers were considered to be those whose mother tongue was the only one spoken in their household but who might possess knowledge of foreign languages acquired through language instruction. All monolingual speakers were either born in St. Petersburg or arrived there as young children from other regions of Russia or the countries of the former Soviet Union.

Since the participants' proficiency in Russian was not explicitly tested, we calculated an individual innovation ratio score as a language performance measure for each monolingual and HS in the sample. More specifically, we manually annotated the innovations in the domain of morphology and lexicon in the oral and written productions according to the classification of the Russian Learner Corpus (Rakhilina et al. 2016). Then, we calculated the individual innovation ratio score by dividing the overall number of innovations

[5] Those HSs who were not able to write in Cyrillic were allowed to write in Latin script.

produced by one speaker by the overall number of words per narration produced by the same speaker. The individual innovation ratio was taken as an indicator of the language performance.

The individual innovation ratio scores were compared among the speakers of the three groups. First, the HSs in the US and in Germany had comparable innovation ratio scores (W = 41, p = 0.382).[6] Next, the HSs in the US showed innovation scores that were significantly higher than the scores of the monolingual speakers (W = 61, p = 0.001). Finally, the HSs in Germany also showed innovation scores that were significantly higher than the scores of the monolingual speakers (t = 2.513, df = 14, p = 0.02). To sum up, both HS groups demonstrated similar innovation ratio scores, while monolingual speakers showed significantly fewer innovations compared to the HSs.

4.3. Data Analysis

This section provides information about the basic annotation layers of the RUEG corpus and additional manual syntactic annotations[7] of the data. In the RUEG corpus, all the collected data were anonymized, and each participant was assigned a code, as shown in Table 3 on the following page. The data were annotated on the so-called diplomatic layer, dipl. This layer provides word-level segmentation without taking the orthographic rules of standard Russian into account. The spoken data were annotated in Praat (Boersma 2001), while the written texts were automatically transferred to the dipl layer without any modifications. Some texts were written in the Latin alphabet. Each document was transcribed by two student assistants on the basis of the dual control principle in order to minimize typing errors and increase the accuracy. The sample used in the present study was drawn from the dipl layer of the RUEG-RU corpus.

[6] The Unpaired Wilcoxon Rank Sum Test was chosen for the comparisons between HS groups and between the HSs in the US and monolingual speakers since the data of the HSs in the US are not normally distributed. A two-tailed unpaired t-test was chosen for the comparison between the HSs in Germany and monolingual speakers since the data of these two speaker groups are normally distributed.

[7] Semi-automatic syntax annotation of the RUEG-RU corpus in SUD framework (Gerdes et al. 2018, 2019; Gerdes and Kahane 2016) is still in progress.

Table 3. Participants' codes

RU, DE, US	country of elicitation: Russia, Germany, US
mo, bi	mode of language acquisition: monolingual, bilingual
01–99	participants' number
F, M	gender: female, male
R	heritage language or, for monolinguals, their only language: Russian
fs	mode of elicitation: formal spoken
is	mode of elicitation: informal spoken
fw	mode of elicitation: formal written
iw	mode of elicitation: informal written
R	language of elicitation: Russian

After the data had been exported from the corpus, they were enriched with manual annotations for syntactic functions. The annotations were carried out by two PhD students of linguistics who are L1 speakers of Russian. The annotations were crosschecked by the two aforedescribed students. First, the utterances were tagged for verb type, word order, and clause type (main vs. embedded clause),[8] as shown in the following examples:

(5) Main clause

…sobaka:	(-) è:	pobežala	za	mjačom.
dog	uh	run	after	ball

'…the dog uh ran after the ball.' (USbi63MR_fsR)

(6) Embedded clause

Kogda	sobaka	uvidela	futbol'nyj	mjač…
when	dog	saw	soccer	ball

'When the dog saw the soccer ball…' (USbi68FR_fwR)

Verbless clauses and clauses with no object were excluded from further analysis. Second, for the declarative clauses with OV or VO orders, the object

[8] See for details Zuban et al. (2021).

realization (nominal or pronominal) was annotated. The final dataset consists of 96 documents and 1,010 data points with either a finite or non-finite verb and at least one pronominal or nominal object (direct and/or oblique) each. The examples below illustrate typical OV and VO clauses in the dataset:

(7) …i ženščina eë vzjala na ruki.
and woman her took on arms
O, pronominal **V**
'…and the woman took her (it) in her arms.' (USbi66FR_fwR)

(8) …on poterjal mjač.
he lost ball
V **O, nominal**
'…he lost the ball.' (DEbi58MR_iwR)

5. Results

As mentioned in §4.3, 1,010 OV and VO clauses were included in the dataset for further analysis. Table 4 provides absolute and relative frequencies of OV and VO orders aggregated by the country of elicitation. The numerical values provide an overall impression that monolingual speakers and the HSs in Germany behave similarly to each other, while the HSs in the US show a higher rate of VO orders compared to the two aforementioned groups.

Table 4. Distribution of OV/VO word orders

Country of elicitation	OV, abs. (rel.)	VO, abs. (rel.)	SUM
Germany	122 (30.5%)	278 (69.5%)	400
Russia	95 (33.2%)	191 (66.8%)	286
US	79 (24.4%)	245 (75.6%)	324

Tables 5 and 6 illustrate the distribution of OV and VO orders grouped according to the object realization, i.e., nominal and pronominal objects. It can be seen (Table 5) that all speaker groups prefer VO orders with nominal objects. As for the pronominal objects (Table 6), although all speaker groups prefer OV orders, this preference is less pronounced in the heritage group from the US than in the two other groups. This descriptive analysis gives an

impression that the first observed differences between some speaker groups might also be statistically significant.

Table 5. Distribution of OV/VO with nominal objects

Country of elicitation	OV, abs. (rel.)	VO, abs. (rel.)	SUM
Germany	43 (14%)	264 (86%)	307
Russia	51 (22.9%)	172 (77.1%)	223
US	41 (15.9%)	217 (84.1%)	258

Table 6. Distribution of OV/VO with pronominal objects

Country of elicitation	OV, abs. (rel.)	VO, abs. (rel.)	SUM
Germany	79 (84.9%)	14 (15.1%)	93
Russia	44 (69.8%)	19 (30.2%)	63
US	38 (57.6%)	28 (42.4%)	66

Finally, tables 7 and 8 show the distribution of OV and VO orders in main and embedded clauses across different speaker groups. It can be seen that all speaker groups prefer VO orders over the OV orders across different clause types. Besides, all speaker groups demonstrate an increase of the VO orders in the embedded clauses, and this increase is particularly substantial for the HSs in the US.

Table 7. Distribution of OV/VO in main clauses

Country of elicitation	OV, abs. (rel.)	VO, abs. (rel.)	SUM
Germany	86 (32.3%)	180 (67.7%)	266
Russia	81 (35.7%)	146 (64.3%)	227
US	73 (29.2%)	177 (70.8%)	250

Furthermore, we applied a generalized linear mixed-effects model which has the following advantages over the descriptive analysis. First, statistical

modeling can account for the speaker's random effect. Besides, such modeling allows investigation of multiple fixed effects. Furthermore, the generalized linear mixed-effect model is especially attractive when the data of each participant are not independent since the model can deal with clusters of dependent data by including the preferences of every individual speaker in the analysis (Winter 2019).

Table 8. Distribution of OV/VO in embedded clauses

Country of elicitation	OV, abs. (rel.)	VO, abs. (rel.)	SUM
Germany	36 (26.9%)	98 (73.1%)	134
Russia	14 (23.7%)	45 (76.3%)	59
US	6 (8.1%)	68 (91.9%)	74

The following predictors were included in the model: *country of elicitation* (Germany, US, and Russia), *clause type* (main vs. embedded clause), *object realization* (nominal vs. pronominal object), *formality* (formal vs. informal), and *mode* (spoken vs. written).[9] The model included the interactions of *country of elicitation* with the fixed effects of *clause type* and *object realization* since these interactions were directly relevant for our research questions and hypotheses.[10] The fixed effects were contrast-coded using sum contrast coding (either –0.5 or 0.5). For post-hoc tests of interactions (pairwise Tukey test), R's (R Core Team 2021) emmeans package (Lenth 2023) was used. *Participants* were treated as a random effect (SD = 0.5283). Furthermore, the main factor of interest, *country of elicitation,* had three levels (Germany, US, Russia), one of which was set as a reference level, namely Russia, since the present study compares the heritage varieties of Russian with the monolingual one.

The model revealed two main effects and one interaction (for the model summary and confidence intervals, see Table 9 and Table 10 in the Appendix). First, there was the main effect of *clause type*: the likelihood of producing an OV word order was lower in the embedded clauses than in the main clauses (see Figure 2 in the Appendix). In the embedded clauses, VO orders were almost always preferred by all speaker groups ($z = –2.57$, $p < 0.05$). Second, there

[9] The last two predictors, *formality* and *mode,* were included in the model due to the experimental setup that explicitly tested the role of formality and mode in the linguistic behavior of heritage and monolingual speakers.

[10] The predictors *formality* and *mode* could not be included as interactions with *country of elicitation* since the model failed to converge.

was the main effect of *object realization*: all speaker groups were more likely to produce an OV order with the pronominal object, while the speakers' preference shifted towards a VO order when the object was realized by a noun ($z = -6.33$, $p < 0.001$) (see Figure 3 in the Appendix).

Furthermore, there was a significant two-way interaction between *country of elicitation* and *object realization*. For the interaction of country and object realization, the Tukey multiple comparison test was run. The Tukey's adjustment revealed the following meaningful significant results. First, as reported earlier, all speaker groups preferred an OV order with the pronominal object, and they preferred a VO order with a nominal object (HSs in the US: $z = -6.13$, $p < 0.001$; HSs in Germany: $z = -10.38$, $p < 0.001$; monolingual speakers: $z = -6.33$, $p < 0.001$). Second, the HSs in Germany produced an OV order with the pronominal object significantly more frequently than the HSs in the US, and the former group produced a VO order with a pronominal object significantly less frequently than the latter group ($z = 3.81$, $p < 0.01$). Note that the HSs of both groups did not differ from the monolingual speakers regarding their preference for the pronominal object with the OV orders (HSs in the US/monolinguals: $z = 1.6$, $p = 0.6013$; HSs in Germany/monolinguals: $z = -2.27$, $p = 0.2037$) (see Figure 1).

To sum up, *clause type* and *object realization* were found to play an important role in the OV/VO choice by all speaker groups in a similar manner. Besides, the HSs in Germany differed from the HSs in the US by producing OV orders with the pronominal object more frequently. Finally, both HS groups were similar to the monolingual speakers regarding their preference for the OV/VO orders and the effect of *clause type* and *object realization*. The predictors *formality* and *mode* turned out to be insignificant.

6. Discussion

The present study investigated the OV/VO choice in heritage and monolingual Russian, using semi-spontaneous, manually annotated data and applying a generalized linear mixed-effects model. In the following, the research questions and hypotheses will be addressed in detail.

The first research question focused on the possible influence of the majority English on the choice of OV/VO orders in productions of the HSs in the US. Hypothesis 1 stated that HSs would produce VO orders more frequently than monolingual speakers regardless of the object realization in both main and embedded clauses.

The data refute Hypothesis 1 since there were no significant differences between the HSs in the US and the monolingual speakers. Apparently, the numerical differences that arose in the descriptive analysis of the data (e.g., the overall distribution of OV/VO orders between the HSs in the US and the monolingual speakers or the potential differences between these two speaker

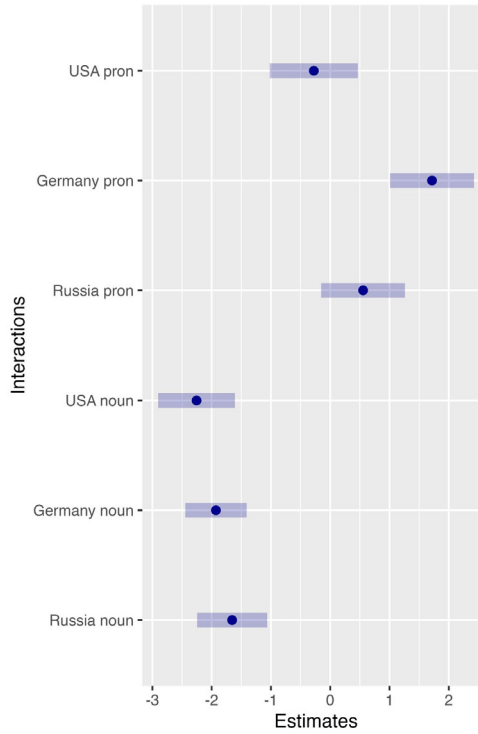

Figure 1. Likelihood of an OV word order:
Interaction of country of elicitation and object realization

groups in the embedded clauses or with the pronominal objects) turned out to be insignificant when speakers' variation was taken into account. This result contradicts those of the previous studies (e.g., Isurin 2005; Kisselev 2019; Laleko and Dubinina 2018; Polinsky 2006; and Zuban et al. 2021).

The differences between the abovementioned research and the results of this study can be explained with reference to the experimental and statistical methodology.

First, the annotation schemes in the abovementioned studies differ from the annotation scheme used in the present study. The study by Isurin (2005) includes the trivalent orders, namely SVO, SOV, SVX, SXV, VSX, XVS, and XSV, as well as two bivalent orders, SV and VS. The (X)OV(X) and (X)VO(X) orders are not mentioned, and it remains unclear whether they did not appear in the data or whether they were not annotated. Polinsky (2006: 238) does not provide any details about the annotation of word order but explicitly points out that the overuse of SVO needs to be considered cautiously due to the difficulties of annotating word order (the speech of HSs typically contains

many pauses). Laleko and Dubinina (2018) annotated clauses with a verbal predicate and at least one of the following constituents: subject, object, or a verbal modifier (adverbial). Clauses with the following word order were considered to be canonical: (S) (Modifier) V (DO) (IO) (Modifier). Clauses with at least one constituent aligned in a different way relative to the verbal predicate were considered to be non-canonical (Laleko and Dubinina 2018: 202). Thus, canonical clauses could contain, for instance, orders without any object (e.g., SV or only with a Modifier V). The annotation scheme in the study by Kisselev (2019) is based on six trivalent word order patterns, SVO, SOV, VSO, VOS, OSV, and OVS, and two bivalent word orders, SV and VS. Thus, OV and VO clauses without subjects were not included in the analysis. This is also true for the study by Zuban et al. (2021), where only trivalent word orders with realized subjects were accounted for.

All in all, the annotation schemes described above are only partially equivalent to the annotation scheme used in the present study (i.e., OV and VO orders with the overtly expressed object and other constituents being optional). Thus, the reoccurring result of the previous studies about the prevalence of the SVO order does not necessarily imply the prevalence of the VO order over OV.

Apart from the annotation scheme, the data analysis might have led to the discrepancies between the current study and the previous studies. Specifically, the present study, contrary to Kisselev (2019), Laleko and Dubinina (2018), and Zuban et al. (2021),[11] used statistical modeling that took into account the individual speaker variation. Corpus studies that look at the *inter-group* variation without the *within-group* variation have been criticized since, in this case, the independence assumption is violated, making the results vulnerable to the Type I error (false positive) (Baayen et al. 2008; Brezina and Meyerhoff 2014; and Winter 2011, 2019).

The second research question focused on the possible influence of the majority German on the choice of OV/VO orders in productions of the HSs in Germany. Hypothesis 2 stated that HSs would produce OV orders more frequently than monolingual speakers in embedded clauses with nominal objects.

Our data seem to provide evidence against Hypothesis 2 since there were generally no significant differences between the HSs in Germany and the monolingual speakers. Specifically, the results of the study showed that in embedded clauses HSs in Germany favor VO word orders with nominal objects, similar to monolingual speakers. Thus, no support for the possible influence of the ML German was found. These results contradict those of Brehmer and Usanova (2015). Although the current study and Brehmer and Usanova

[11] This does not apply to Isurin (2005) and Polinsky (2006) since the former study investigates one single participant while the latter study is a survey article.

(2015) used a similar approach to annotate the data, these two studies differed regarding the statistical method. Brehmer and Usanova (2015), contrary to this study, did not take the individual speaker variation into consideration.

Given these two aspects, namely different annotation schemes and statistical approaches, the results of the present study are not surprising, since neither of these two aspects appears to be fully comparable to those of the previous studies. Thus, the results of the present investigation are novel, inviting replications and further research using the advanced methodology.

In the following, the results will be discussed in terms of their general impact. First and foremost, our results are in line with previous research on monolingual Russian confirming a strong tendency for a postverbal linearization of nominal objects and a preverbal linearization of pronominal objects (Kallestinova 2007). Importantly, we found that this trend is not only true for the monolingual speakers but for the HSs as well. Apparently, this feature remains stable even under language contact. These results clearly speak against transfer effects from MLs.

Although we did not aim at testing the Interface Hypothesis (Sorace 2011), we will briefly discuss possible consequences for it. Our results can be interpreted in two ways. They may be taken to speak in favor of considering head-directionality within the VP as being the decisive factor for the distribution of OV/VO patterns in Russian, i.e., under this view, the overall OV/VO distribution is rather determined by core-grammatical (internal) principles. Alternatively, if one maintains the view that the overall OV/VO distribution is determined by information structural conditions, these results together with some previous studies may raise some questions about the predictive power of the Interface Hypothesis since external interfaces are claimed to be prone to variability under language contact (e.g., Slabakova et al. 2012 on L2 learners; Jin et al. 2022; Leal et al. 2014; and Méndez et al. 2015 on HSs). It might be the case that not all external interfaces are similarly challenging for HSs. It is also possible that the distinction between the internal and external interface is not a completely reliable predictor for the linguistic behavior of bilingual speakers (both L2 learners and HSs).

Furthermore, our results indicate that OV/VO choice is clause-type sensitive. Specifically, the likelihood of producing an OV order was lower in the embedded clauses than in the main clauses. This result suggests that both heritage and monolingual speakers tend to use (S)VO more frequently in embedded clauses than in main clauses. In fact, a closer look at embedded clauses with VO orders reveals that almost all the word orders (210 out of 211)[12] are of (S)VO(O) type. For monolingual Russian, it was already stated that word order in embedded clauses is less varied compared to main clauses, probably due to the differences of the two clause types in terms of their

[12] One clause uttered by a monolingual speaker has a VOS word order.

discourse structure (Bailyn 2012 and Zuban et al. 2021). Our findings provide support for this claim. However, one of the limitations of the present study is that the data were not controlled for information-structural appropriateness. Since word order patterns of HSs of Russian were not always found to be contextually appropriate (Brehmer and Usanova 2015; Laleko and Dubinina 2018), we assume that it might also be true for our data sample. Moreover, the authors are aware of the fact that the dataset used in the present study is not exhaustive in its size and homogeneity. Once the syntactic annotation in the RUEG-RU corpus is complete, further studies have to be conducted to reproduce the results reported in the present paper on a larger number of observations. Also, other relevant factors that can influence word order in Russian, such as intonation, animacy of the object, or grammatical weight of the constituents, were not considered in the present research. Future studies should look at these factors and their possible interplay since they may give new additional insights into the OV/VO choice.

The results of our study can be explained by the participants' high engagement with Russian. According to the self-reports, the HSs participating in our study actively use Russian in their daily life from birth (communication with caregivers, extra activities in Russian, media use in Russian) and almost all HSs can write in the Cyrillic script. This is especially remarkable if one compares the HSs' profiles from the present study with those in the studies carried out by Isurin (2005) and Polinsky (2006), where the participants were barely involved in any kind of activities using their HL. Finally, the HSs of both groups were found to differ from each other regarding their preference for OV/VO orders with the pronominal object. These results indicate that the HSs of Russian residing in different countries may behave differently from each other. It is difficult to pinpoint the exact reason behind these results. It is possible that the differences in the speaker communities or some other extra-linguistic factors contributed to the differences between the two HS groups.

To sum up, the HSs in the US and Germany were similar to the monolingual speakers of Russian regarding their choice of OV/VO orders. On one hand, these results call for a new evaluation of the previous studies and their implications for the possible cross-linguistic effects from the MLs. We believe that language transfer may arise in some particular populations of HSs (presumably those that are not extensively engaged in the practices involved in using Russian, e.g., speaking, listening, overhearing, responding, code-switching, etc.), while for other HS populations that are highly engaged with their HL, the transfer is not necessarily at play. On the other hand, we suggest that the factors that are relevant for word order choice in monolingual Russian (i.e., object realization and clause type) are also relevant for heritage Russian in the US and Germany to a similar extent.

Sources

Wiese, Heike, Artemis Alexiadou, Shanley Allen, et al. (2021) RUEG Corpus. Available at: https://zenodo.org/record/5808870.

References

Aboh, Enoch Oladé. (2015) *The emergence of hybrid grammars: Language contact and change*. Cambridge, UK: Cambridge University Press.

Anstatt, Tanja. (2011) "Russisch in der zweiten Generation. Zur Sprachsituation von Jugendlichen aus russischsprachigen Familien in Deutschland". Ludwig M. Eichinger, Albrecht Plewnia, and Melanie Steinle, eds. *Sprache und Integration: Über Mehrsprachigkeit und Migration*. Tübingen: A. Francke Verlag, 101–28.

Aref'ev, Aleksandr L. (2012) *Russkij jazyk na rubeže XX–XXI vekov*. Moscow: Centr social'nogo prognozirovanija i marketinga.

Baayen, R. Harald, Douglas J. Davidson, and Douglas M. Bates. (2008) "Mixed-effects modeling with crossed random effects for subjects and items". Special issue: Emerging data analysis, *Journal of memory and language* 59(4): 390–412. DOI 10.1016/j.jml.2007.12.005.

Bailyn, John. (1995) *A configurational approach to Russian free word order.* PhD dissertation, Cornell University.

————. (2004) "Generalized inversion". *Natural language & linguistic theory* 22(1): 1–50.

————. (2012) *The syntax of Russian*. Cambridge, UK: Cambridge University Press.

Baker, Mark. (2001) *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. (2015) "Fitting linear mixed-effects models using lme4". *Journal of statistical software* 67(1): 1–48. DOI 10.18637/jss.v067.i01.

Benmamoun, Elabbas, Silvina Montrul, and Maria Polinsky. (2013) "Heritage languages and their speakers: Opportunities and challenges for linguistics". *Theoretical linguistics* 39(3–4): 129–81. DOI 10.1515/tl-2013-0009.

Birner, Betty, and Gregory Ward. (1998) *Information status and noncanonical word order in English*. Amsterdam: John Benjamins.

Boersma, Paul. (2001) "Praat, a system for doing phonetics by computer". *Glot international* 5(9–10): 341–45.

Brehmer, Bernhard. (2007) "Sprechen Sie Qwelja? Formen und Folgen russischdeutscher Zweisprachigkeit in Deutschland". Tanja Anstatt, ed. *Mehrsprachigkeit Bei Kindern und Erwachsenen. Erwerb, Formen, Förderung.* Tübingen: Narr, 163–86.

Brehmer, Bernhard, and Grit Mehlhorn. (2015) "Russisch als Herkunftssprache in Deutschland. Ein holistischer Ansatz zur Erforschung des Potenzials von Herkunftssprachen". *Zeitschrift für Fremdsprachenforschung* 26(1): 83–121.

Brehmer, Bernhard, and Irina Usanova. (2015) "Let's fix it? Cross-linguistic influence in word order patterns of Russian heritage speakers in Germany". Hagen Peukert, ed. *Transfer effects in multilingual language development.* Amsterdam: John Benjamins, 159–86.

Brezina, Vaclav, and Miriam Meyerhoff. (2014) "Significant or random. A critical review of sociolinguistic generalisations based on large corpora". *International journal of corpus linguistics* 19(1): 1–28.

Broekhuis, Hans. (2006) "The universal base hypothesis: VO or OV?" Jeroen van de Weijer and Bettelou Los, eds. *Linguistics in the Netherlands 2006.* Amsterdam: John Benjamins, 28–39.

Cho, Grace, Fay Shin, and Stephen Krashen. (2004) "What do we know about heritage languages? What do we need to know about them?" *Multicultural education* 11(4): 23–26.

Cuza, Alejandro. (2012) "Crosslinguistic influence at the syntax proper: Interrogative subject-verb inversion in heritage Spanish". *International journal of bilingualism* 17(1): 71–96. DOI 10.1177/1367006911432619.

Cuza, Alejandro, and Joshua Frank. (2011) "Transfer effects at the syntax-semantics interface: The case of double *que* questions in heritage Spanish". *Heritage language journal* 8(1): 66–89.

Dubinina, Irina, and Maria Polinsky. (2013) "Russian in the USA". Michael Moser and Maria Polinsky, eds. *Slavic languages in migration*. Münster: Lit Verlag.

Eppler, Eva. (1999) "Word order in German-English mixed discourse". *UCL working papers in linguistics* 11: 285–308.

Fenyvesi, Anna. (2005) "Hungarian in the United States". Anna Fenyvesi, ed. *Hungarian language contact outside Hungary: Studies in Hungarian as a minority language.* Amsterdam: John Benjamins, 265–318.

Gagarina, Natalia. (2014) "Die Erstsprache bei Mehrsprachigen im Migrationskontext". Solveig Chilla and Stefanie Haberzettl, eds. *Mehrsprachigkeit*. Vienna: Elsevier, 19–37.

————. (2017) "Monolingualer und bilingualer Erstspracherwerb des Russischen: Ein Überblick". Kai Witzlack-Makarevich and Nadja Wulff, eds. *Handbuch des Russischen in Deutschland: Migration – Mehrsprachigkeit – Spracherwerb.* Berlin: Frank & Timme, 393–410.

Gagarina, Natalia, and Annegret Klassert. (2018) "Input dominance and development of home language in Russian-German bilinguals". *Frontiers in communication* 3: 40. DOI 10.3389/fcomm.2018.00040.

Gagarina, Natalia, Maria Martynova, Irina Sekerina, and Luka Szucsich. (2020) "From first verbs to adulthood: Aspect in heritage Russian in the US and Germany". Evgenij Vasil'evič Golovko, Elena Viktorovna Gorbova, Petr Aleksandrovič Kočarov, Viktor Samuilovič Xrakovskij, and Oksana Jur'evna Čujkova, eds. *Vzaimodejstvie aspekta so smežnymi kategorijami. Materialy VII Meždunarodnoj konferencii Komissii po aspektologii Meždunarodnogo komiteta slavistov.* Petersburg: RGPU im. A. I. Gercena, 112–17.

Gärtner, Hans-Martin. (2000) "Are there v2 relative clauses in German?" *The journal of comparative Germanic linguistics* 3(2): 97–141.

Gerdes, Kim, Bruno Guillaume, Sylvain Kahane, and Guy Perrier. (2018) "SUD or Surface-Syntactic Universal Dependencies: An annotation scheme near-isomorphic to UD". Universal Dependencies Workshop 2018, Brussels, Belgium. Available at: https://hal.inria.fr/hal-01930614. Last accessed 31 October 2024.

————. (2019) "Improving Surface-Syntactic Universal Dependencies (SUD): Surface-syntactic relations and deep syntactic features". TLT 2019—18th International Workshop on Treebanks and Linguistic Theories, Paris, France. Available at: https://hal.inria.fr/hal-02266003. Last accessed 31 October 2024.

Gerdes, Kim, and Sylvain Kahane. (2016) "Dependency annotation choices: Assessing theoretical and practical issues of universal dependencies". Annemarie Friedrich and Katrin Tomanek, eds. *Proceedings of the 10th Linguistic Annotation Workshop held in conjunction with ACL 2016 (LAW-X 2016).* Berlin: Association for Computational Linguistics, 131–40. DOI 10.18653/v1/W16-1715.

Guijarro-Fuentes, Pedro, and Katrin Schmitz. (2015) "The nature and nurture of heritage language acquisition". *Lingua* 164: 239–50. DOI 10.1016/j.lingua.2015.05.008.

Haider, Hubert. (2012) *The syntax of German*. Cambridge, UK: Cambridge University Press.

————. (2015) "Head directionality - in syntax and morphology". Antonio Fábregas, Jaume Mateu, and Michael Putnam, eds. *Contemporary linguistic parameters*. London: Bloomsbury Academic, 73–97.

Haider, Hubert, and Luka Szucsich. (2022) "Slavic languages–'SVO' languages without SVO qualities?" *Theoretical linguistics* 48(1–2): 1–39.

Hamann, Katharina, Kai Witzlack-Makarevich, and Nadja Wulff. (2019) "Russian in Germany". Arto Mustajoki, Ekaterina Protassova, and Maria Yelenevskaya, eds. *The soft power of the Russian language: Pluricentricity, politics and policies*. London: Routledge, 163–74. DOI 10.4324/9780429061110-14.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. (2008) "Simultaneous inference in general parametric models". *Biometrical journal* 50(3): 346–63.

Isurin, Ludmila. (2005) "Cross linguistic transfer in word order: Evidence from L1 forgetting and L2 acquisition". James Cohen, Kara T. McAlister, Kellie Rolstad, and Jeff MacSwan, eds. *Proceedings of the 4th International Symposium on Bilingualism*. Somerville, MA: Cascadilla Press, 1115–30. Available at: http://www.cascadilla.com/isb4.html. Last accessed 31 October 2024.

————. (2011) *Russian diaspora: Culture, identity, and language change*. Berlin, New York: De Gruyter Mouton. DOI 10.1515/9781934078457.

Isurin, Ludmila, and Tanya Ivanova-Sullivan. (2008) "Lost in between: The case of Russian heritage speakers". *Heritage language journal* 6(1): 72–104.

Jasinskaja, Katja, and Radek Šimík. (forthcoming) "Slavonic free word order". Jan Fellerer and Neil Bermel, eds. *The Oxford guide to Slavonic languages*. Oxford, UK: Oxford University Press.

Jin, Jing, Sihui Echo Ke, and John Chi-Kin Lee. (2022) "Language interfaces in adult heritage language acquisition: A study on encoding of nominal reference in Mandarin Chinese as a heritage language". *Frontiers in psychology* 12. DOI 10.3389/fpsyg.2021.790102.

Johannessen, Janne Bondi, and Signe Laake. (2015) "On two myths of the Norwegian language in America". Janne Bondi Johannessen and Joseph C. Salmons, eds. *Germanic heritage languages in North America*. Amsterdam: John Benjamins, 299–322. Available at: https://www.jbe-platform.com/content/books/9789027268198-silv.18.14joh.

Junghanns, Uwe, and Gerhild Zybatow. (1995) "Syntax and information structure of Russian clauses". Wayles Browne, Ewa Dornisch, Natasha Kondrashova, and Draga Zec, eds. *Annual Workshop on Formal Approaches to Slavic Linguistics: The Cornell Meeting*. Ann Arbor: Michigan Slavic Publications, 289–319.

Kallestinova, Elena D. (2007) *Aspects of word order in Russian*. Iowa City: University of Iowa.

Kayne, Richard. (1994) *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kempen, Gerard, and Karin Harbusch. (2019) "Mutual attraction between high-frequency verbs and clause types with finite verbs in early positions: Corpus evidence from spoken English, Dutch, and German". *Language, cognition and neuroscience* 34(9): 1140–51. DOI 10.1080/23273798.2019.1642498.

King, Tracy H. (1995) *Configuring topic and focus in Russian*. Stanford, CA: CLSI Publications.

Kisselev, Olesya. (2019) "Word order patterns in the writing of heritage and second language learners of Russian". *Heritage language journal* 69: 149–74.

Kroch, Anthony, and Ann Taylor. (2000) "Verb-object order in early middle English". *Diachronic syntax: Models and mechanisms* 132: 163.

Laleko, Oksana. (2010) *The syntax-pragmatics interface in language loss: Covert restructuring of aspect in heritage Russian*. PhD dissertation, University of Minnesota.

————. (2011) "Restructuring of verbal aspect in heritage Russian: Beyond lexicalization". *International journal of language studies* 5(3): 13–26.

————. (2013) "Assessing heritage language vitality: Russian in the United States". *Heritage language journal* 10(3): 89–102.

————. (2015) "From privative to equipollent: Incipient changes in the aspectual system of heritage Russian". Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau, and Maria Yastrebova, eds. *Slavic grammar from a formal perspective*. Frankfurt am Main: Peter Lang, 273–86.

————. (2018) "What is difficult about grammatical gender? Evidence from heritage Russian". *Journal of language contact* 11(2): 233–67. DOI 10.1163/19552629-01102004.

————. (2019) "Resolving indeterminacy in gender agreement: Comparing heritage speakers and L2 learners of Russian". *Heritage language journal* 16(2): 151–82.

Laleko, Oksana, and Irina Dubinina. (2018) "Word order production in heritage Russian: Perspectives from linguistics and pedagogy". Susan Bauckus and Susan Kresin, eds. *Connecting across languages and cultures: A heritage language Festschrift in honor of Olga Kagan*. Bloomington, IN: Slavica Publishers, 191–215.

Laleko, Oksana, and Maria Polinsky. (2017) "Silence is difficult: On missing elements in bilingual grammars". *Zeitschrift für Sprachwissenschaft* 36(1): 135–63. DOI 10.1515/zfs-2017-0007.

Leal, Tania, Jason Rothman, and Roumyana Slabakova. (2014) "A rare structure at the syntax-discourse interface: Heritage and Spanish-dominant native speakers weigh in". *Language acquisition* 21(4): 411–29.

Lenth, Russell V. (2023) "Emmeans: Estimated marginal means, aka least-squares means" [R package version 1.8.6]. Available at: https://CRAN.R-project.org/package=emmeans. Last accessed 31 October 2024.

Lobanova, Anna. (2011) "The role of prominence scales for the disambiguation of grammatical functions in Russian". *Russian linguistics* 35(1): 125–42.

Luchkina, Tatiana, and Jennifer S. Cole. (2016) "Structural and referent-based effects on prosodic expression in Russian". *Phonetica* 73(3–4): 279–313.

Méndez, Tania L., Jason Rothman, and Roumyana Slabakova. (2015) "Discourse-sensitive clitic-doubled dislocations in heritage Spanish". *Lingua* 155: 85–97.

Montrul, Silvina. (2015) *The acquisition of heritage languages*. Cambridge, UK: Cambridge University Press. DOI 10.1017/CBO9781139030502.

O'Grady, William, Hye-Young Kwak, On-Soon Lee, and Miseon Lee. (2011) "An emergentist perspective on heritage language acquisition". *Studies in second language acquisition* 33(2): 223–45. Available at: http://www.jstor.org/stable/44486002.

Pintzuk, Susan, and Ann Taylor. (2006) "The loss of OV order in the history of English". Ans van Kemenade and Bettelou Los, eds. *The handbook of the history of English*. Carlton, Australia: Blackwell Publishing, 249–78.

Polinsky, Maria. (2006) "Incomplete acquisition: American Russian". *Journal of slavic linguistics* 14(2): 191–262.

————. (2008a) "Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization". *Heritage language journal* 6(1): 40–71.

————. (2008b) "Without aspect". G. G. Corbett and M. Noonan, eds. *Case and grammatical relations*. Amsterdam: John Benjamins, 263–82. Available at: https://www.benjamins.com/catalog/tsl.81.

————. (2011) "Reanalysis in adult heritage language: A case for attrition". *Studies in second language acquisition* 33: 305–28.

————. (2015) "Heritage languages and their speakers: State of the field, challenges, perspectives for future work, and methodologies". *Zeitschrift für Fremdsprachwissenschaft* 26: 7–27.

————. (2018) *Heritage languages and their speakers*. Cambridge, UK: Cambridge University Press. DOI 10.1017/9781107252349.

R Core Team. (2021) "R: A language and environment for statistical computing". Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/. Last accessed 31 October 2024.

Rakhilina, Ekaterina V., Anastasia Vyrenkova, Elmira Mustakimova, Alina Ladygina, and Ivan Smirnov. (2016) "Building a learner corpus for Russian". Elena Volodina, Gintarė Grigonytė, Ildikó Pilán, Kristina Nilsson Björkenstam, and Lars Borin, eds. *Proceedings of the joint workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition*. Umeå, Sweden: LiU Electronic Press, 66–75.

Romanova, Natalia. (2008) "Mechanisms of verbal morphology processing in heritage speakers of Russian". *Heritage language journal* 6(1): 105–26.

Rothman, Jason. (2009) "Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages". *International journal of bilingualism* 13(2): 155–63. DOI 10.1177/1367006909339814.

Slabakova, Roumyana, Paula Kempchinsky, and Jason Rothman. (2012) "Clitic-doubled left dislocation and focus fronting in L2 Spanish: A case of successful acquisition at the syntax-discourse interface". *Second language research* 28(3): 319–43.

Slioussar, Natalia. (2007) *Grammar and information structure. A study with reference to Russian*. Utrecht: LOT.

————. (2011) "Processing of a free word order language: The role of syntax and context". *Journal of psycholinguistic research* 40(4): 291–306.

Sorace, Antonella. (2011) "Pinning down the concept of 'interface' in bilingualism". *Linguistic approaches to bilingualism* 1(1): 1–33.

Statistisches Bundesamt. (2019a) "Bevölkerung in Privathaushalten nach Migrationshintergrund im weiteren Sinn nach Geburtsstaat in Staatengruppen. Ergebnisse des Mikrozensus". Available at: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/migrationshintergrund-staatsangehoerigkeit-staatengruppen.html. Last accessed 23 April 2020.

————. (2019b) "Pressemitteilung Nr. 314 vom 21. August 2019: Jede vierte Person in Deutschland hatte 2018 einen Migrationshintergrund". Available at: https://www.destatis.de/DE/Presse/Pressemitteilungen/2019/08/PD19_314_12511.html. Last accessed 23 April 2020.

Švedova, Natal'ja Jul'evna, ed. (1980, 2005) *Russkaja grammatika*. Moscow: Institut russkogo jazyka im. V.V. Vinogradova RAN.

Topaj, Nathalie. (2018) *Grammatical and pragmatic use of referential expressions in picture-based narratives of bilingual and monolingual children in Russian and German*. PhD dissertation, Humboldt-Universität zu Berlin.

United Nations. (2019) "International migrant stock 2019: Department of Economic and Social Affairs, Population Division". Available at: www.unmigration.org. Last accessed 29 April 2020.

Uriagereka, Juan. (2007) "Clarifying the notion 'parameter'". *Biolinguistics* 1: 99–113.

US Census Bureau. (2015) "Detailed languages spoken at home and ability to speak English for the population 5 years and over: 2009–2013". Available at: https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html. Last accessed 1 July 2021.

Valdés, Guadalupe. (2005) "Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized?" *The modern language journal* 89(3): 410–26. DOI 10.1111/j.1540-4781.2005.00314.x.

Wegener, Heide. (1993) "Weil – das hat schon seinen Grund. Zur Verbstellung in Kausalsätzen mit weil im gegenwärtigen Deutsch". *Deutsche sprache* 21: 289–305.

Wiese, Heike. (2020) "Language situations: A method for capturing variation within speakers' repertoires". Yoshiyuki Asahi, ed. *Methods in dialectology*. Volume 15. Bamberg: Bamberg Studies in English Linguistics.

Winter, Bodo. (2011) "Pseudoreplication in phonetic research". *ICPhS XVII*: 2137–40.

Zuban, Yulia, Maria Martynova, Sabine Zerbian, Luka Szucsich, and Natalia Gagarina. (2021) "Word order in heritage Russian: Clause type and majority language matter". *Russian linguistics* 45(3): 253–81.

Appendix

Table 9. Model summary

Predictors	Estimates	SE	z value	*p*
(Intercept)	–0.55024	0.28320	–1.943	5.202450e–02
Germany	0.44474	0.38519	1.155	2.482609e–01
USA	–0.71568	0.42193	–1.696	8.984753e–02
Nominal Objects	–2.21013	0.34942	–6.325	2.529034e–10
Emb. Clauses	–1.04441	0.40618	–2.571	1.013147e–02
Spoken Mode	0.29567	0.17171	1.722	8.507501e–02
Formal Setting	–0.03266	0.17127	–0.191	8.487528e–01
Germany : Nominal Objects	–1.43580	0.49541	–2.898	3.752886e–03
USA : Nominal Objects	0.23169	0.47586	0.487	6.263370e–01
Germany : Emb. Clauses	0.97298	0.51926	1.874	6.096140e–02
USA : Emb. Clauses	–0.32743	0.62513	–0.524	6.004306e–01
Observations	1010			
Marginal R^2 / Conditional R^2	0.317 / 0.371			

Table 10. Confidence intervals

Predictors	2.5%	97.5%
(Intercept)	–1.10531185	0.004823809
Germany	–0.31022657	1.199701204
USA	–1.54265041	0.111289172
Nominal Objects	–2.89497793	–1.525288231
Emb. Clauses	–1.84051030	–0.248316023
Spoken Mode	–0.04086513	0.632211324
Formal Setting	–0.36835491	0.303027532
Germany : Nominal Objects	–2.40677639	–0.464817764
USA : Nominal Objects	–0.70098285	1.164368365
Germany : Emb. Clauses	–0.04475457	1.990717871
USA : Emb. Clauses	–1.55267288	0.897806089

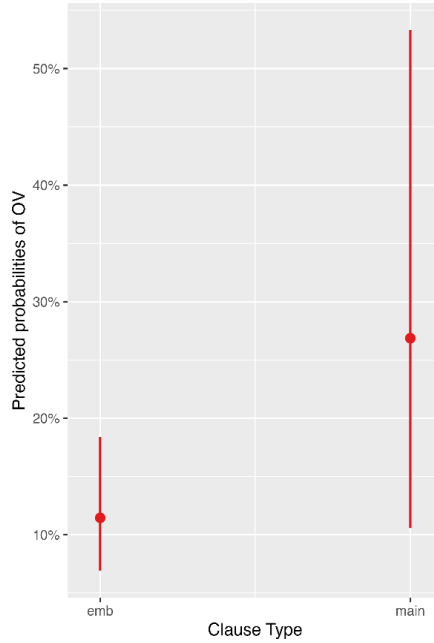

Figure 2. Likelihood of an OV word order in main and embedded clauses

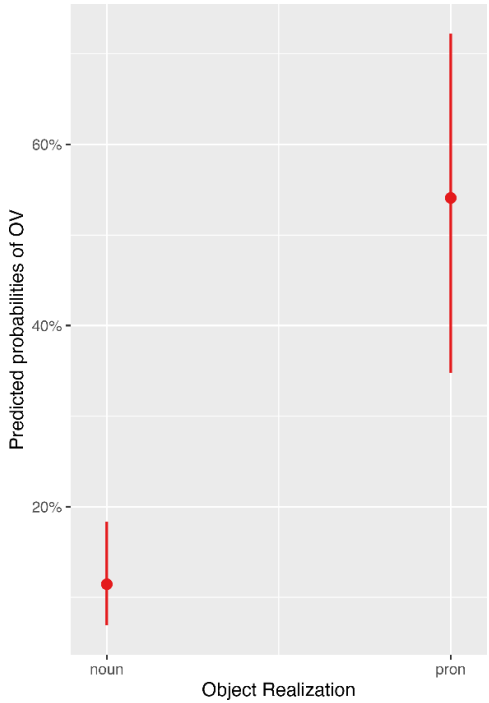

Figure 3. Likelihood of an OV word order with nominal and pronominal objects

Maria Martynova
Humboldt University Berlin
Institute of Slavic Studies
and Hungarology
Berlin, Germany
maria.martynova@gmx.com

Yulia Zuban
University of Stuttgart
Institute of Linguistics
Division of English Linguistics
Stuttgart, Germany
zubanyulia@gmail.com

Luka Szucsich
Humboldt University Berlin
Berlin, Germany
luka.szucsich@hu-berlin.de

Natalia Gagarina
Leibniz Center General Linguistics
Berlin, Germany
gagarina@leibniz-zas.de