2019

Word Order Patterns in the Writing of Heritage and Second Language Learners of Russian

Olesya Kisselev

Word Order Patterns in the Writing of Heritage and Second Language Learners of Russian

Olesya Kisselev

1. Introduction

Word Order (WO) variability is an important feature of the Russian language. Appropriate use of WO patterns makes a Russian text meaningful and coherent and has larger implications for the grammaticality of sentences and the ability of the language user to interpret and convey the meaning of the utterance. In the words of the late Olga Kagan, "every learner and teacher of Russian would agree that acquisition of native-like WO is one of the most challenging hurdles on the path to the higher levels of language performance" (Kagan and Dillion 2004, 89). Despite this widely shared opinion, little is known about the development of WO variability in Russian interlanguage, both in the case of mainstream foreign learners (L2) of Russian and in the case of speakers of Russian as a heritage language (HL) (Laleko and Dubinina 2018).

The purpose of the current study is to address the gap in the existing research literature and to explore WO variation in written Russian learner data as well as to discuss the implications for pedagogy. The study investigates the use of WO patterns from the developmental perspective by comparing the use of WO patterns by students at the intermediate level of language proficiency and the use of WO patterns by students of more advanced language proficiency. Additionally, the study compares the use of WO patterns in the writing of learners from different linguistic backgrounds, L2 and HL. The study is exploratory in nature: with few previous studies addressing WO in Russian learner production, the research aim is formulated broadly as an attempt to describe the patterns of use of different WOs in the writing of learners of Russian and to explore the abilities of the learners to express meanings coded in variable WOs.

2. Russian WO in the light of discourse-pragmatics

The Russian language belongs to the so-called variable WO languages, which—due to their rich inflectional systems—do not have to rely on the order of sentence constituents to mark grammatical functions (see, e.g., Yokoyama 1986; Comrie 1987; Bailyn 2012). Morphological marking can typically help distinguish subjects, objects, and, where available, indirect objects, allowing the constituents to linearize variably. Using this logic, a simple transitive sentence with the proposition can potentially result in six variations of WO: SVO, OVS, SOV, OSV, VSO, and VOS. Despite the availability of all six variations, the actual distribution of the WO patterns appears to be skewed, and Russian NSs show a strong preference for producing some WO patterns over others (Bivon 1971; Kallenstinova 2007), with up to 80% of all Russian sentences realized as SVO. The apparent imbalance in the frequencies of WO patterns is a result of FUNCTIONAL properties of linearization properties.

Largely, variability of WO patterns in Russian is tied to basic discourse functions, which could be described as follows:

1. introducing a NEW TOPIC or referent to the stretch of discourse, usually by asserting the existence of the referent in some "possible world" (Yokoyama 1986, 182);

2. providing ADDITIONAL INFORMATION about the topic or referent that has been introduced earlier (or activated); and

3. providing a STANCE or evaluation of the topic or referent.

Based on these functions, the Russian language is thought to operate with the three basic types of WO patterns: presentational, informational, and expressive (Grenoble 1998; Yanko 2001).

INFORMATIONAL WO is the most frequent and prototypical (basic) of the WOs in Russian (Grenoble 1998, 161). Its main discourse goal, as implied by its name, is to provide additional information on already-known discourse referents or topics, thus developing the discourse further. Consider (1)(b) below. The TOPIC of the sentence (namely, the person named Émma) is the information already known to the listener (as evident in (1)(a)). The new information, i.e., the discipline that Émma studies in college, is the COMMENT on the known topic; it bears the greatest informational load in the sentence.

(1)

a. – *Что изучает Эмма в колледже?*
   'What does Émma study in college?'
b. – *Эмма изучает психологию.*
   Émma studies psychology.

The WO realizing the function of introducing new discourse topics is called PRESENTATIONAL WO (Grenoble 1998, 163). Consider (2)(b) below. The prepositional phrase *в её жизни* 'in her life' is the type of easily identifiable information that follows from the previous sentence that in some ways describes Émma's life (her being at the university). However, the predicate and the subject group ('*появились*' and '*новые друзья*' respectively) are all new information; both bear the information load of the sentence.

(2)

a. – *В прошлом году Эмма поступила в университет.*
   'Last year, Émma started college.'
b. – *В её жизни появились новые друзья.*
   'New friends entered her life.'

In addition to the core sentence constituents (predicate and subject), presentational WO often contains another element, known as the LOCALIZER or the DETERMINANT SENTENCE CONSTITUENT (cf. Russian *детерминант* [Shvedova et al. 1980]). In (2)(b), it is expressed with the phrase *в её жизни* 'in her life'.

An important difference between informational and presentational WO is how this functional distinction is grammaticalized: if the default WO found in the informational sentence type is SV(O), the presentational WO is NORMALLY VS.

The third discourse function is communicated through EXPRESSIVE WO (Grenoble 1998, 161). This type of WO fronts a sentential constituent that bears the greatest informational load, to create either an emphatic or a contrastive reading. In an EMPHATIC sentence, a fronted constituent characterized by an emphatic stress has to introduce new information to the hearer, as in (3)(b); in a CONTRASTIVE sentence, a fronted constituent normally represents identifiable information, as in (4)(b) (Kallestinova 2007).

(3)

a. – *Кто съел все пирожные?*
    'Who ate all the pastries?!'
b. – *Эмма съела все пирожные!*
    'It was Émma who ate all the pastries!'

(4)

a. – *Кто из ваших детей изучает психологию, Федя?*
    'Which one of your kids studies psychology, Fedia?'
b. – *Эмма изучает психологию.*
    'It is Émma who studies psychology.'

It is important to note that the so-called expressive focus is a result of the interaction between WO and intonation. Prosodically, information focus is distinguished by a normal sentential stress (Dyakonova 2004; Kallestinova 2007). In a prototypical sentence that is organized in accordance with the "given first/focus second" principle, focus bears the main prosodic prominence of the sentence (see Chomsky 1971 as cited in Dyakonova 2004, p. 91). Consider the neutral sentence in (2)(a), in which the sentence stress is a normal sentential stress, with a rising contour Low High concluded by a falling tone High Low on the segment denoting the information focus (i.e., *поступила в университет* 'entered the university').

Expressive focus, on the other hand, is characterized by emphatic stress: the intonational contour in (b)(3), for example, is realized by an emphatic focal stress (falling High Low) on the fronted element *Эмма* 'Émma.'

There is a wealth of literature that explicates the role of intonation in the ordering of sentence constituents (see, e.g., Yokoyama 1985, 1986; Yanko 2001; Kiss 1987; Paducheva 2004, 2010). And although prosodic means have a lowered significance in writing compared to speech, in principle, the same parameters apply to written speech, with perhaps greater usage of lexicogrammatical and syntactic means of information highlighting, such as focusing constructions, that highlight fronting of focused material (Callies 2009, 5).

In addition to information-structure consideration, syntactic (or grammatical) WEIGHT has been reported to contribute to positioning of sentential elements. Syntactic weight (sometimes referred to as HEAVINESS)

is understood in terms of the internal structure of a sentence constituent as measured in number of words/syllables and/or morphosyntactic complexity. The preference for LIGHTER constituents to be placed to the left of the HEAVIER ones holds cross-linguistically (Siewierska 1988). The tendency for more complex parts to occur in clause-final position is known as the principle of increasing constituents or "end-weight" principle (Quirk et al. 1972). In accordance with this principle, such light elements as personal, possessive, and demonstrative pronouns, as well as monosyllabic adverbs, are likely to be placed before longer (read, heavier) elements, such as full determiners and determiner phrases (Laleko and Dubinina 2018, 195). Notice that in (5)(b), the light element expressed by the pronoun *её* 'she.ACC' is placed left-ward of the verb, resulting in SOV WO. This WO is a highly preferred (if not the only felicitous) WO in cases where no other heavy constituents (such as obliques) are present along with subject, predicate, and light element. (5)(c) is informationally infelicitous since it places light (and known) elements in the focal position.

(5)
a. – *Вы знаете Надю?*
   'Do you know Nadia?'
b. – *Да, я её знаю.*
   'Yes, I know her.'
c. – **Да, я знаю её.*
   'Yes, I know her.'

In a study that presents evidence from a corpus-based analysis and elicitation experiments, Arnold et al. (2000) reported that both information status of sentence constituents and the syntactic weight of constituents strongly correlate with sentence position. In a different corpus-based study, Wasow (1997) offered an explanation of the end-weight principle in cognitive terms, that it helped the speaker plan the utterance. Light elements can "buy time" in the on-line process of speech production. Although investigations of syntactic weight typically focus on the structural properties of discourse rather than its information structure, the principle of weight-end is related to the principle of "focus second." As observed by Arnold et al. (2003), elements that have been introduced in the previous stretch of discourse can now be referred to using deictic

markers, typically anaphoric pronouns, which are short (i.e., light in grammatical weight). The elements that introduce new information are more likely to be heavy (34). At the same time, the relationship between syntactic weight and sentence position, especially when it comes to the Russian language, is relatively understudied and many questions pertaining to the variability in placement of light constituents remain unexplored.

3. WO in Russian learner language

One of the earlier studies examining WO in the production by American learners of Russian was Thompson (1996). Thompson regarded WO errors in the speech of L2 Russian speakers as a case of discrepancy between the English fixed WO, which itself marks grammatical relations in the sentence and the Russian flexible WO, which necessitates marking the grammatical relations with the help of morphological markers. Similarly, she explained the lack of the VS WO in the speech of American learners of Russian as being due to the absence of such a structure in English. In her paper, Thompson provided a number of examples of erroneous sentences, but her analysis, unfortunately, did not provide a comprehensive or even numeric picture of the results; she also did not consider any of the discourse-pragmatic qualities of different Russian WOs. The Russian researcher Khavronina (2005–2006), surveying WO errors in the speech of learners of Russian from various L1s, concluded that WO is difficult for all learners of Russian, regardless of L1 and level of proficiency (128). Khavronina suggested that the errors in WO stem from the learners' lack of awareness of the "sentence bipartition" (i.e., the differences in the informational load as given or old information). Although the paper is descriptive in nature and lacks any numerical data or analysis, it, too, gives additional credence to the general observations regarding the difficulty learners of Russian face when dealing with Russian WO patterns.

The somewhat more numerous studies of WO in HL Russian have yielded mixed results. One of the earlier studies, by Polinsky (2006), suggested a relatively strong retention of the VS WO in the oral production of HL speakers regardless of the general level of language proficiency (237). However, this optimistic conclusion was not universally upheld in other studies. Kagan and Dillon (2004), for example, found a

significant reduction in the use of VS patterns in their HL data. Having examined a small corpus of elicited oral narratives (n=18) produced by HL students (age of emigration from 0 to 10), the authors found a mere five clauses displaying the VS pattern in all 18 HL narratives (compared to 11 sentences in just one NS story). Most importantly, all VS WOs in the HL data were a type of cliché *жили-были* (cf. English "There once were X") used at the beginning of stories.

Kagan and Dillon's (2004) findings were supported in a later study by Isurin and Ivanova-Sullivan (2008), who found an equally small number of VS instances (n=8) in narratives elicited from six HL speakers. The authors reported that this number of VS clauses in the HL data amounted to 2.1% of all patterns, compared to VS clauses amounting to 6% of patterns found in the monolingual Russian NS data (99). Interestingly, an even smaller percentage, only 1%, was found in comparable L2 data. The authors argued that there exists a link between the occurrence of the VS pattern and the length of exposure to Russian among both HLs and L2 learners (100); however, Isurin and Ivanova-Sullivan recognized that the relatively small participant sample in their study may have hampered the ability to generalize their findings about noncanonical sentences over the population of Russian L2 and HL learners.

Even if the general observation regarding the reduction in WO flexibility and overreliance on SVO WO in HL (and likely L2) language is correct, the "specific manifestations of such general reduction in WO variation have not been discussed at length" (Laleko and Dubinina 2018, 197). Most importantly, the variability (or lack of thereof) in WO was not discussed in those papers in terms of discourse-pragmatics, the underlying reason for the existence of such variability.

The most comprehensive account of WO as a product of discourse-pragmatic requirements is presented in Laleko and Dubinina (2018). Unlike previous studies, Laleko and Dubinina (2018) found a considerable proportion of HL clauses to fit the noncanonical category, i.e., the non-SV(O) pattern (22%), although this percentage was statistically smaller than that for the NS data (32%). The authors further examined the types of noncanonical patterns, namely, INVERSION (presentational WO in my terminology) and DISLOCATION (expressive WO in my terminology, such as OSV), as well as the CONTEXTUAL APPROPRIATENESS of the chosen WOs. Again, both HLs and NSs aligned in their preferences for dislocated

patterns versus inversed patterns; however, the HL data contained a considerable proportion of informationally infelicitous WOs. The canonical WOs contained the least amount of errors: only 3% of all canonical clauses were categorized as informationally infelicitous. There was a greater proportion of infelicitous VS constructions (9%) and an even greater percentage of infelicitous clauses with dislocation (30%). The authors concluded that the HLs employ "different strategies" in the use of the two types of WO patterns; more specifically, they use dislocation more frequently overall but "nevertheless fall short of using dislocation in a target-like way, possibly as a result of non-target-like principles governing its occurrence" (205). The VS pattern, on the other hand, is used less frequently but far more appropriately, which "indicates a more target-like control of principles that govern its use" (205).

The detailed account on the use of WO in HL Russian presented in Laleko and Dubinina (2018) reconciled some of the controversial findings in the previous literature. More importantly, by teasing apart the complexities of WO use (or underuse) in bilingual production, the study underscored the necessity of further exploration of WO patterns in their relation to the discourse-pragmatic distinctions they realize.

The overall conclusions regarding the studies of Russian L2 and HL learners' use of WO patterns—however few studies there are on this topic at the moment—seem to align with the results and generalizations made in research on other language pairs. A growing number of studies (e.g., Schachter and Rutherford 1979; Rutherford 1983; Von Stutterheim, Carroll and Klein 2003; Green et al. 2000; Bohnacker and Rosén 2008; Callies 2009; Jackson and Ruf 2017) all come to the overall conclusion that L2 speakers exhibit non-native preferences for ordering of sentence constituents not only at the sentential level (at the level of syntax) but more broadly in the domain of information organization. At this level, a learner has to figure out not only possible alternatives and their functional properties but also the constraints on the use of possible WO patterns. Transfer of principles from the dominant language to the L2 results in texts that are "unidiomatic" and "not fully cohesive from the perspective of a native speaker" (Bohnacker and Rosén 2008, 534).

The aim of the present study is to contribute to a growing body of research on WO variation in Russian learner data. By examining the discourse-pragmatic functions of Russian WO in the data produced by HL

and L2 speakers of Russian, I set out to evaluate whether naturalistic and early exposure to the target language results in a more target-like use of WO patterns. By comparing the use of WO patterns between students at a lower level of language proficiency and those of more advanced language proficiency, I explored whether (and how) the ability to manipulate WO patterns grows with proficiency.

4. Methods

4.1. Participants and data collection

The data in this study are drawn from a corpus of essays drawn from the annual American Council of Teachers of Russian National Post-Secondary Russian Essay Contest (hereafter, Contest). Hundreds of students, representing 30 to 40 U.S. universities and colleges, voluntarily participate in the contest each year. The students are grouped according to their approximate exposure to the language. For L2 learners of Russian, the grouping is determined by the number of instruction hours they have received by the time of the contest (e.g., fewer than 100 hours for the lowest proficiency group, 100–250 hours for the next proficiency level, and so on). For HL learners, the level is determined by the approximation of naturalistic exposure to the language (level 1 includes HL learners of Russian who were either born in the United States or had emigrated before the age of six and had no formal instruction in a Russian-speaking country prior to enrolling in a college-level course.

The parameter used for distinguishing the HL groups is clearly less than perfect. It does not take into account many factors that contribute to the overall proficiency of a HL speaker, such as the amount of language exposure at home or experience with semi-formal instruction in Russian through after-school activities. Nonetheless, the number of authors in the sub-corpus allows us to reasonably expect that the possible contamination of data has been well mitigated.

In this study, I focus on the more proficient L2 learners of Russian, since coherent and cohesive discourse (and, thus, a clear need to mark the discourse-pragmatic distinctions) emerges in L2 learners at Intermediate-Mid level on the ACTFL scale, a level that could be expected of the students after more than 250 hours of instructed Russian language study. Direct comparison of language proficiency levels in L2 and HL groups is difficult to make (although HL speakers normally place at Intermediate

level in speaking), and I do not assume similar writing abilities in the HL and non-HL students. The purpose of this study is to investigate the different groups of writers in their own right.

For the comparison, a set of essays (*n* ranging from 21 to 23) were randomly selected from the two higher-proficiency L2 groups (the "FL3 group", i.e., the L2 learners with approximately 250–400 hours of instructional experience, and the "FL4" group, i.e., the L2 learners with more than 400 hours of instructional experience) and the HL group, including only the learners who were either born in an English-speaking country or emigrated before the onset of schooling. To have a comparable reference corpus that would be of a similar genre and created in similar experimental conditions, I collected essays on the same topic from 17 Russian NSs, young adults who were either living in Russia or who recently arrived in the United States. The descriptive statistics of the four sets of data are presented in Table 1.

Table 1. Number of word tokens and sentences in HL, FL3, FL4,
and NS sub-corpora

Group	Measure	Total	Mean	SD	Min.	Max.
HL	Tokens	5,348	232.3	133.151	97	605
(N = 23)	Sentences	454	19.74	10.230	8	45
FL3	Tokens	5,298	238.00	97.350	85	419
(N = 22)	Sentences	478	21.95	10.472	9	55
FL4	Tokens	6,241	296.48	111.558	95	576
(N = 21)	Sentences	504	24.00	8.803	10	40
NS	Tokens	5,309	295.59	82.129	165	484
(N = 17)	Sentences	410	22.94	8.112	12	44

The texts obtained through the Contest are a unique set of data representing dozens of language programs across the country (and, therefore, various instructional approaches) as well as various proficiency levels and language-learning backgrounds. Additionally, the data are collected in similar settings with the same constraints

and affordances for all participants. The fact that the topic is the same across programs and levels also allows for more meaningful comparisons among the groups. Although the Contest participants are not instructed to write in any specific genre, I found that in response to the prompt *What is a friend?*, most reacted with a short expository essay with elements of narration (e.g., autobiographical events). Since WO patterns—just as other linguistic categories—are found to be distributed differently in texts of different communicative purposes (Turner 2006; McAnallen 2009), analyzing texts collected through similar procedures and for similar purposes makes the between-group comparisons more meaningful.

4.2. Data preparation and annotation

All selected texts were separated into clauses for further coding and analysis. The scope of the study was narrowed to include only the more canonical type of clauses, namely, a declarative indicative clause with an overt nominal or pronominal subject in the nominative case[1] and a predicate (S_{nom}). The advantage of considering more canonical sentences is that they allow us to assume with a greater degree of confidence that the actual WO pattern produced by the learner is not prompted by a difficulty with a rarer or more specific syntactic construction and that that WO pattern is the result of the interplay of the two requisite forces—the concerns of discourse-pragmatics and the rules of sentential grammar. Each clause was then coded for WO pattern and discourse-pragmatic felicitousness of the WO pattern. It is important to note that the infelicitous-use category here includes only errors in the ordering of major constituents; all other structural issues (such as misplacement of adverbs or particles, or morphosyntactic errors) are disregarded for purposes of this study. Secondly, each clause was examined with regard to its specific intended discourse function. Coding for discourse-pragmatic felicitousness was executed by two NSs of Russian with training in linguistics; the few discrepancies were discussed and resolved. A clause was deemed infelicitous only when the chosen WO clearly resulted in a breach in the flow of discourse.

---

[1] A very small category of subjects also included in the final dataset consists of subjects expressed through quantifier+NP (e.g., two friends, many people), which some linguists consider noncanonical subjects.

5. Results

5.1. Quantitative results

The numeric results for coding the declarative S_{nom} clauses for WO patterns are presented in Table 2. The table reports the first six WO patterns in the order of NS "preference" for specific WO realizations found in Kallestinova (2007). The two additional WO patterns (SV and VS) were tallied separately and are reported directly below the order into which they are traditionally subsumed; thus, the SV clauses can be seen below the SVO numbers, and the VS clauses can be seen directly below the OVS numbers. The data are presented in this way in order to allow for a more straightforward comparison with past research that addresses WO distribution patterns (specifically, Kallestinova 2007).

Table 2. Raw and prorated frequencies of WO patterns in the HL, FL3, FL4 and NS sub-corpora

WO patterns	HL	FL3	FL4	NS
N. of S_{nom} clauses	592	622	636	385
SVO	210 (35%)	267 (43%)	294 (46%)	161 (42%)
SV	252 (43%)	269 (43%)	237 (37%)	134 (35%)
OVS	42 (7%)	42 (7%)	50 (8%)	45 (11.7%)
VS	6 (1%)	5 (1%)	20 (3%)	21 (5.5%)
SOV	73 (12%)	34 (5%)	28 (4.5%)	16 (4%)
VSO	0	0	1 (0.15%)	2 (0.5%)
VOS	0	1 (0.1%)	0	3 (1%)
OSV	9 (1.5%)	4 (0.5%)	6 (1%)	3 (1%)

As demonstrated in Table 2, the relative proportions of the WOs in the NS sub-corpus follow the order of preference for various WOs reported in the previous studies.

Importantly, the types of clauses produced by the L2 learners in both the FL3 and the FL4 groups follow the same distributional pattern,

with some differences in the percentages of these distributions. More specifically, the FL3 and the FL4 learners' SVO and SV WOs are the most frequent, similar to the NS controls; however, the percentage of the SV(O) WO is higher in the L2 data: the SVO and SV clauses combined account for 84% and 86% of the WO patterns found in the FL4 and FL3 data respectively. The percentage of the (O)VS clauses in the L2 data is smaller than those found in the NS data: 11% for the FL4 learners and 7% for the FL3 learners.

The HL learners can also be said to prefer the SV WO to a greater extent than the NNs: in addition to SVO and SV clauses (which together account for 78% of WO patterns combined), the HLs produced the largest number of SOV clauses of all four groups (an additional 12% of all data, a number that stands in contrast to the rest of the writers). The (O)VS clauses in the HL sub-corpus are almost as infrequent as in the FL3 data and amount to 8% of all clauses. A chi-square analysis showed that the differences in the proportion of different WO patterns across the four data sets are statistically significant ($\chi^2 = 82.388$, p < .0001)[2]. Pairwise chi-square analyses further revealed differences between the HL and FL3 groups ($\chi^2 = 32.49$, p < .0001), the HL and FL4 groups ($\chi^2 = 34.485$, p < .0001), the HL and NS groups ($\chi^2 = 33.677$, p < .0001), and the FL4 and FL3 groups ($\chi^2 = 14.957$, p < .01). While the difference between the NS and FL3 groups is significant ($\chi^2 = 23.87$, p < .001), the difference between the higher-proficiency FL4 group and the NSs was found to be not significant ($\chi^2 = 6.471$, p = .263).

Another important perspective on the differences in the WO usage between the four groups emerged from the comparison of all patterns in which the subject occupies pre-verbal position (SV-pattern) to all patterns in which the subject occupies the post-verbal position (VS-pattern). As expected the VS clauses are less numerous in my data, aligning with the general observations about Russian WO. However, all three learner groups produced notably fewer VS clauses than the NSs. It appears that the learners relied more heavily on the canonical WOs. The percentages of SV clauses formed a cline, with the FL3 group producing the largest amount of SV clauses (92%), followed by the HL learners (91%), followed by the FL4 group (89%).

---

[2] A note on the chi-square analysis: the counts for the VSO and VOS clauses were excluded from the analysis since the numbers are very small (from zero to 3) and a chi-square test does not allow zeros in its calculations.

Table 3. Raw and prorated frequencies of SV WO patterns and VS WO patterns in the HL, FL3, FL4, and NS sub-corpora

Types of clauses	HL	FL3	FL4	NS
N. of Snom clauses	592	622	636	385
SVO	210	267	294	161
SV	252	269	237	134
SOV	73	34	28	16
OSV	9	4	6	3
The sum of all SV-pattern	544	574	565	314
clauses	**91%**	**92%**	**89%**	**81%**
VOS	0	1	0	3
VS	6	5	20	21
OVS	42	42	50	45
VSO	0	0	1	2
The sum of all VS-pattern	48	48	70	71
clauses	**8%**	**8%**	**11%**	**18%**

Pairwise chi-square analyses demonstrated that these differences in the proportions of SV- vs VS-patterns are significant between the HL and NS groups (χ^2= 23.288, p < .0001), between the FL3 and NS groups (χ^2= 26.247, p < .0001), between the FL4 and NS groups (χ^2= 11.071, p < .0001), and between the FL4 and FL3 groups (χ^2= 4.039, p = .04). The HL writers occupy a middle ground between the lower- and the higher-level L2 learners: they are neither statistically different from the lower-lever L2 learners, nor statistically different from the higher-level L2 learners. The differences between the HL and either the FL3, or the FL4 group do not reach statistical significance (χ^2= 2.99, p = .0834 and χ^2= .064, p = .8002 respectively).

Overall, the results indicate that all groups of learners utilized all the same WOs that are available in standard Russian. More importantly, they utilized the various patterns in the same order of preference established in the NS data (in the current study, as well as in previous research). In the same manner as the NSs, the learners produced a significantly larger proportion of SV-patterns than the VS-patterns. However, the statistical tests reveal that all three learner groups relied a lot more on the canonical SV-patterns than the NSs did, underutilizing the noncanonical WO. The statistical test also revealed differences between the learner groups: the lower-level L2 learners were more likely to produce a canonical SV-pattern than the higher-level L2 learners. The HL learners are not statistically different from either L2 group.

These patterns indicate that the learners underutilize WO as a linguistic tool and may miss opportunities to use variable WO to signal important pragmatic meanings.

5.2. Qualitative analysis

To account for the apparent underuse of the particular WO types, I further examined each WO pattern with regard to its discourse-pragmatic function.

First, each clause extracted for analysis in this paper was marked as pragmatically felicitous or infelicitous. Numerically, only a small proportion of clauses in the learner data represented a clear misuse of the chosen WO (these cases are listed as INFELICITOUS USE in Table 4 below).

However, even a small number of clear misusage of the appropriate WO coupled with the "missed opportunities," i.e., contexts in which an alternative WO would have been preferred, reflected in the significant differences of the different types of WO patterns reported above, suggest that learners experience difficulties in choosing appropriate WO to achieve the specific communicative goals. The analysis that follows provide a further exploration of various WO patterns in learner data, with the exception of VOS and VSO patterns, which do not appear in the learner data and are too rare in the NS data to arrive at any conclusions.

Table 4. Infelicitous clauses across types of WO patterns in HL, FL3, FL4, and NS sub-corpora

	HL	FL3	FL4	NS
N. of S_{nom} clauses	592	622	636	385
SV(O) clauses	462	536	531	295
Infelicitous use of SVO and SV WOs	8	13	4	n/a
SOV	73	34	28	16
Infelicitous use of SOV WO	7	5	5	n/a
OSV	9	4	6	3
Infelicitous use of OSV WO	3	4	2	n/a
(O)VS clauses	47	47	70	46
Infelicitous use of (O)VS WOs	1	1	3	n/a
VOS	0	0	0	3
VSO	0	0	1	2
Infelicitous use of VOS/VSO WOs	0	0	1	n/a

5.2.1. SV-pattern clauses

As follows from the numerical results, all three learner groups are apt users of the canonical SV(O) WO. The function of providing additional information on the known topics is well mastered by all learners. Nonetheless, infelicitous SVO clauses (on top of general patterns of SVO overuse) do appear in the learner data, with the lower-proficiency group performing somewhat worse than the other two groups. Both the FL3 group and the HL group (albeit in fewer instances) produced SV(O) clauses instead of the obligatory VS presentational WO constructions, as in the example below, where the learner produced an SVO clause instead of an obligatory VS (three such errors were found in the HL data and six in the FL3 data).

(6)
a. *Иногда, когда я дома мне всё скучно.*
   Sometimes, when I am home, I am bored.

b. *Но потом мой друг пришёл.*
　　But then my friend came (comes?).
(FL3_2264)

The higher-level L2 writers did not appear to use SVO instead of the required presentational VS.

In addition to its primary function, the SV-pattern also appears in learner data in its expressive function, a type of clause in which new information is placed at the beginning of a sentence to create either an emphasis or a contrastive reading. However, the lower-level proficiency L2 group cannot be said to have mastered this pattern: all 18 SV(O) clauses with NEW subjects were categorized as infelicitous. The HL learners and the FL4 learner produced fewer SV(O) clauses with NEW subjects (n=13 and n=9 respectively) with four such clauses in each group being marked as infelicitous. Thus, in regards to expressive SV(O), the HL speakers may have a slight advantage over instructed L2 learners; however, the level of proficiency in L2 speakers clearly plays a role in the ability to produce an SVO expressive clause.

Another type of expressive SV-pattern is OSV. In this construction, the leftward position of the object [3] may be motivated by an intention to place emphasis on the object.

Although the OSV pattern is rather infrequent in my data, it was used by all four groups of writers. The NSs used three OSV clauses; while the HL learners and the FL4 learners produced a slightly higher number of the OSV clauses than the NS controls did: nine OSV clauses by the HL learners and six OSV clauses by the FL4 learners. Examples of successful realization of OSV patterns are shown below in (7)(b) and (8)(b):

(7)
a. *До того времени я считал их друзьями,*
b. *А Джерри я считал человеком, как отец.*
　　'Before then I considered them my friends, but Jerry, I considered to be father-like to me.'
(FL4_5716)

---

[3] The position of an object or an adjunct in the pre-verbal position is termed DISLOCATION in many theoretical and empirical works on WO (see Bailyn 2012), including the paper by Laleko and Dubinina (2018) reviewed above.

(8)

a. *Он встречался с девушками,*

b. *но любовь он никогда не знал и не понимал.*

  'He was dating girls, but love, he never knew nor understood.'

(HL_2223)

Although not all of these clauses were felicitous like (7)(b) and (8) (b), the percentage of errors is low: two infelicitous OSV clauses were found in the HL sub-corpus and one in the FL4 sub-corpus. The FL3 learners, on the other hand, seemed to be less apt in dealing with the OSV pattern; they attempted fewer OSV clauses (n=4), and all four were found to be infelicitous.

5.2.2. *VS and OVS WOs*

As demonstrated above, all groups of learners used comparable numbers of VS-patterns, all three underusing this pattern in comparison to the NSs. At the same time, all learners appeared to use the (O)VS WO in its appropriate discourse function, i.e., introducing new discourse topics. Some important differences in the use of presentational WO by different learner groups also persist. For instance, the FL3 writers' use of localizers in this structure is limited; there are only three clauses expressing presentational WO that have a temporal or spatial determinant. And although sentences like (9) are grammatically correct and WO-felicitous, the use of a localizer (such as *в России* 'in Russia') could have helped to anchor the following new information.

(9)

Есть поговорка, «Доброе слово и кошке приятно.»

'There is a saying: "Even a cat will appreciate a kind word."'

(FL3_9438)

HL and FL4 learners, on the other hand, utilized more presentational WO constructions, and they appeared to use more localizers, which help establish the shared context between the writer and the reader. However, unlike the HL data, the FL4 data contain multiple examples of positioning locatives at the end of the clauses, in accordance with the preferred information structure of the English sentence (cf. (10)(b) produced by a learner to (11)(b) Standard Russian).

(10)

a. *Вопрос о том, что такое друг, сложнее, но в основном я думаю,*

 'The question of What is a friend is more complex [than this], but overall I think.'

b. *что есть культурная разница <u>в понятии слова «друг»</u>.*
<div align="center">*localizer*</div>

 'that there is a cultural difference in the notion of the word "friend".'

(FL4_0158)

(11)

a. *Вопрос о том, что такое друг, сложнее, но в основном я думаю,*

b. *что <u>в понятии слова «друг»</u> есть культурная разница.*
      localizer

(Standard Russian)

5.2.3. SOV WO

The SOV WO is a particular variation of the canonical SVO WO, in which the order of the subject and the verb in relation to one another remains canonical (SV), but the object is placed in the pre-verbal position, rendering the whole construction a noncanonical WO in Russian.[4] The difference between the SVO and SOV patterns is that in the latter both S and O are known information (whereas in SVO an object can – but not necessarily – mark new information). Most importantly the object in the SOV clause is highly likely to be realized by a syntactically light element.

As shown in Table 2, the SOV pattern is well represented in all four sub-corpora. To summarize, both in absolute numbers and in percentages, the SOV pattern appeared on a cline, with the NSs producing the least amount of SOV clauses (n=16, 4% of all WO patterns), followed by the FL4 learners (n=28, 4% of all WOs), followed by the FL3 learners (n=34, 5.5%), and with the HL learners producing the largest number of SOV patterns (n=73, 12%), a substantial proportion of all WO patterns the HL learners produced.[5]

---

[4] In structuralist approaches, such a variation on the canonical WO is referred to as DISLOCATION, as opposed to INVERSION, whereby the order of subject and verb is realized as VS.
[5] These numbers do not account for temporal, locative or manner obliques, which are considered light elements and tend to "move" leftward. In this analysis, I considered only elements that were coded as obligatory arguments and were given the code "Object."

(12)
Иногда я его ненавижу.
          light Object
'Sometimes I hate him.'
(HL_7649)

All groups of learners appeared to adhere to "end-weight" (Quirk et al. 1972) principle by placing light elements in the preverbal position. Yet, the HL learners' relative preference for utilizing the SOV pattern is striking. One explanation is that errors in object placement contribute to the overall large tallies; yet, the number of such errors is relatively small (n=3), with one HL speaker producing two of them. This observed HL preference for leftward placement of light constituents aligns with the observations made in Laleko and Dubinina (2018), who found grammatical weight to emerge "as a strong predictor of leftward movement in the heritage language, compared to the baseline" (208). To provide a satisfactory explanation of the observed trend, future studies of syntactic weight in heritage languages are needed.

6. Discussion

The first tangible and important finding of this study is the fact that all learner groups produced all six grammatically possible WO variations and that they produced them on the scale of NS preference established in earlier studies (Bivon 1971; Kallestinova 2007) as well as in this work. The fact that the types of clauses produced by the HL learners and both L2 groups follow the same distributional pattern as the clauses produced by the NSs is noteworthy: it shows that learners of Russian (at least at intermediate and higher proficiency levels) have an overall understanding of the availability of WO patterns and their distributional patterns. More importantly, from the perspective of discourse-pragmatic functions of Russian WO, the learners appeared to be able to distinguish the three basic discourse functions and, overall, appeared to have a good grasp on how the underlying discourse principles are realized in WO. The learners fulfilled the function of providing additional information on known topics by using the SV(O) WOs. They introduced new discourse topics or new referents by employing the (O)VS WO. The learners were also found to use patterns that front some sentence constituents to create an emphatic reading of the utterances.

However, the proportions of different WO patterns across the four data sets were found to differ significantly. The most relevant perspective on the differences emerged from the comparison of all patterns in which the subject occupies pre-verbal position (SV-pattern) to all patterns in which the subject occupies the post-verbal position (VS-pattern). The statistical tests reveal that all three groups of learners relied a lot more on the canonical SV-patterns than the Russian NSs, while at the same time underutilizing the noncanonical WOs. The percentages of SV-pattern clauses formed a sort of cline, with the FL3 group producing the largest amount of SV clauses (92.2%), followed by the HL learners (91.8%), followed by the FL4 group (88.8%). The difference between the FL3 and FL4 group was found to be statistically significant. Evidently, the use of SV(O) becomes more target-like—at least numerically—as language proficiency increases. This conclusion supports the observation of Isurin and Ivanova-Sullivan (2008), who attempted to explain the conflicting results of their study by implicating language proficiency as a factor in the ability to produce variable WO.

Variability of learner proficiency likewise contributed to the type of infelicitous WOs produced by the three groups. For instance, the FL3 learners (and to a lesser extent the HL learners) were found to use the canonical SV(O) WO when the context required the presentational WO (this tendency, of course, is also reflected in descriptive numeric analysis, which demonstrated learners' overreliance on SV-type patterns). This error is more pronounced in the data of the lower-level L2 learners (the percentage of such errors in the FL3 data is 4.8% compared to 3.8% in the HL data).

Proficiency level differentiates the two L2 groups in their use of presentational WO. Not only did the FL3 learners produce fewer presentational VS patterns, but their use of these constructions was extremely rigid: by omitting localizers, they missed opportunities to establish the topic and shared knowledge space where the new information could be anchored. Higher-level L2 learners were found to use a greater variety of presentational WOs, and their use of these constructions was contextually more appropriate and grammatically correct. These learners were closer to NSs in terms of numbers (although still statistically lagging behind), and an overwhelming majority of the (O)VS structures in the FL4 data were structurally sound and

informationally felicitous. Many examples showed that the more advanced learners used variable localizers and a variety of existential verbs with added semantic meanings to fit the context. Clearly, language proficiency plays a role in the use of presentational WO; at the same time, early exposure to language, may also be advantageous for this construction, as the HL learners in this study were found to masterfully use the localizer+VS constructions.

It appears that all learner groups experienced difficulties with the less frequent types of clauses, specifically, clauses with fronted referents. Although all learners exhibited some understanding of the fact that SV(O) WO can realize an expressive function, they produced a large proportion of informationally infelicitous and/or structurally problematic clauses of this type.

A similar picture emerged from the analysis of the OSV and OVS clauses, in which the object is fronted. Object-fronting, which offers an opportunity to add emphasis to the proposition and/or create cohesion between the two clauses, was avoided by the FL3 learners, and the few clauses (n=4) that the FL3 learners attempted were all classified as informationally infelicitous. Higher-level L2 learners and, even more so, the HL learners produced more object-fronted WOs. However, because these clauses require manipulation of the syntactic structure of the sentence, the learners often produced somewhat infelicitous or structurally deviant sentences. It is likely that the discourse function of the expressive WO exists on the conceptual plane; however, when it comes to choosing the appropriate linguistic form (including WO), the learners experience difficulties.

Even though it appeared that the L2 learners were improving their use of WO, the fact that even advanced learners significantly underused the variability of patterns indicates that they continued to miss opportunities to produce more nuanced and more coherent discourse. This is likely a result of instructional history. WO is rarely discussed in Russian language textbooks, and the topic is at best provided a few cursory remarks. None of the textbooks more frequently used in the United States include a functional explanation of WO variation. The case of HL speakers in my study shows that relying on exposure (or implicit learning) when it comes to WO does not guarantee development and explicit instruction of this topic is in order.

The ability to comprehend and create pragmatically appropriate discourse in Russian is dependent on understanding the underlying principles of variability of Russian WO patterns. The examples of a functional approach to teaching Russian WO do exist. The Russian textbook "Word order in Russian sentences" (Krylova and Khavronina 1976) is known to be used in study-abroad programs. Although the effort of Krylova and Khavronina is laudable (if not entirely unquestionable), integrating a stand-alone book intended for advanced learners of Russian into a regular beginner to intermediate level syllabus is unfeasible, since most examples and activities in the book employ more advanced lexicon and syntax. I believe that discussion of WO and the discourse-pragmatic principles that underlie WO should be dispersed throughout the curricula, beginning in the first semester when "basic" structures such as *У меня есть X* 'I have X' and *Там есть X* 'There is X' are first introduced. WO should be regularly revisited as more complex lexicogrammatical structures are introduced to the learners (such as *В этой статье рассматриваются вопросы*, 'The article focuses on such issues as…'). In the absence of such an integrated approach, instructors are unlikely to explicitly deal with pragmatic errors that stem from infelicitous WOs. Thus, the augmentations of the teaching resources do not need to entail a complete overhaul of teaching curricula; rather they should take a form of better and function-based explanations of variable WO that learners are exposed to.

I believe that greater awareness of the importance of information-structural aspects in realizations of linguistic form will also lead to more questions about how it impacts learner language. This, in turn, may spark greater interest in information structure as a topic in Second Language Acquisition studies. One of the reasons why we see so few studies on information structure and, consequently, WO is that the question of what constitutes advanced proficiency in a second language has been and remains focused on mastering grammatical competence, i.e., sentence-level syntax.

Acknowledgements

References

Arnold, Jennifer E., Losongco, Anthony, Wasow, Thomas, and Ginstrom, Ryan. 2000. "Heaviness vs. Newness: The Effects of Structural Complexity and Discourse Status on Constituent Ordering. *Language*, 76 (1): 28-55. doi:10.2307/417392

Bailyn, John F. 2012. *The Syntax of Russian*. Cambridge: Cambridge University Press.

Bivon, Roy. 1971. *Element Order*. Cambridge: Cambridge University Press.

Bohnacker, Ute, and Christina Rosén. 2008. "The Clause-Initial Position in L2 German Declaratives: Transfer of Information Structure." *Studies in Second Language Acquisition* 30 (4): 511–38.

Callies, Marcus. 2009. *Information Highlighting in Advanced Learner English: The Syntax-Pragmatics Interface in Second Language Acquisition*. Amsterdam: John Benjamins.

Chomsky, Noam. 1969. *Deep Structure, Surface Structure, and Semantic Interpretation*. Bloomington, Indiana: Indiana University Linguistics Club.

Comrie, Bernard. 1987. "Russian." In *The World's Major Languages*, edited by B. Comrie, 322–28. New York: Oxford University Press.

Dyakonova, Marina 2004. "Information Structure Development: Evidence Form the Acquisition of Word Order in Russian and English." *Nordlyd* 32 (1): 88–109.

Green, Christofer F., Christopher, Elsie R., & Mei, Jaquelin Lam Kam (2000). "The Incidence and Effects on Coherence of Marked Themes in Interlanguage Texts: A Corpus-based Enquiry. *English for Specific Purposes*, 19 (2): 99-113.

Grenoble, Leonore A. 1998. *Deixis and Information Packaging in Russian Discourse*. Amsterdam: John Benjamins.

Khavronina, Serfaima. 2005–2006. "Pis'mennaia rech' studentov-inostrantsev stochki zreneniia norm russkogo slovoraspolozheniia." *Vestnik RUDN*, series *Russkii i inostrannye iazyki i metodika ix prepodavanija* 1: 126–32.

Isurin, Ludmila, and Tanya Ivanova-Sullivan. 2008. "Lost in Between: The Case of Russian Heritage Speakers." *Heritage Language Journal* 6 (1): 72–104.

Jackson, Carrie, and Helena Ruf. 2017. "The Priming of Word Order in Second Language German." *Applied Psycholinguistics* 38 (2): 315–

45. doi:10.1017/S0142716416000205.

Kagan, Olga, and Kathleen Dillon. 2004. "Heritage Speakers' Potential for High-Level Language Proficiency." *In Advanced Foreign Language Learning: A Challenge to College Programs*, edited by Heidi Byrnes and Hiram H. Maxim, 99–112. Boston, MA: Heinle/Thomson.

Kallestinova, Elena. 2007. "Aspects of Word Order in Russian." PhD diss., University of Iowa. http://ir.uiowa.edu/etd/.

Kiss, Katalin. (1987). "Configurationality in Hungarian, A 'Free' Word Order Language." *Linguistic Inquiry*, 12: 185-215.

Krylova, Olga, and Serafima Khavronina. 1976. *Word Order in Russian Sentences*. Moscow, Russia: Russian Language.

Laleko, Oksana and Irina Dubinina. 2018. "Word Order Production in Heritage Russian: Perspectives from Linguistics and Pedagogy." In *Connecting Across Languages and Cultures: A Heritage Language Festschrift in Honor of Olga Kagan*, edited by Susan Bauckus and Susan Kresin, 191–215. Bloomington, IN: Slavica.

Paducheva, Elena (2004). *Dinamicheskie modeli v semantike leksiki*. Moscow, Russia: Iazyki slavianskoi kul'tury.

Paducheva, Elena (2010). "Subject-Predicate Inversion and its Cognitive Sources." *Russian Linguistics: International Journal for the Study of Russian and Other Slavic Languages*, 34(2), 113-121.

Polinsky, Maria 2006. "Incomplete Acquisition: American Russian." *Journal of Slavic Linguistics* 14 (2): 191–262.

Quirk, Randolph. 1972. *A Grammar of Contemporary English*. London: Longman.

Rutherford, William 1983. "Language Typology and Language Transfer." In *Language Transfer in Language Learning*, edited by Susan M. Gass and Larry Selinker, 358–70. Rowley, MA: Newbury House.

Schachter, Jacquelyn, and William Rutherford. 1979. "Discourse Function and Language Transfer." *Working Papers on Bilingualism Toronto* (19): 1–11.

Shvedova, Natalia, and Vladimir Lopatin (Eds.). (1989). *Kratkaia russkaia grammatika*. Moscow, Russia: Nauka.

Siewierska, Anna 1988. *Word Order Rules*. London: Routledge.

Thompson, Irene. 1996. "Assessing Foreign Language Skills: Data from Russian." *The Modern Language Journal* 80 (1): 47–65.

Von Stutterheim, Christiane, Mary Carroll, and Wolfgang Klein. 2003.

"Two Ways of Construing Complex Temporal Structures." In *Deictic Conceptualization of Time, Space, and Person*, edited by F. Lenz, 97–134. Amsterdam: John Benjamins.

Yanko, Tatiana (2001). *Kommunikativnye strategii russkoj rechi*. Moscow, Russia: Iazyki slavianskoj kul'tury.

Yokoyama, Olga T. 1985. "A Diversified Approach to Russian Word Order." In *Issues in Russian Morphosyntax*, edited by Richard D. Brecht and Michael S., pp. 187-208. Columbus, OH: Slavica.

Yokoyama, Olga T. 1986. *Discourse and Word Order*. Amsterdam: John Benjamins.

Wasow, Thomas (1997). "Remarks on Grammatical Weight." *Language Variation and Change*, 9 (1): 81-105.