Word order and prosody effects in perception of sentence focus in Russian by native speakers and adult L2 learners

TATIANA LUCHKINA[1], TANIA IONIN[2], MARIA GOLDSHTEIN[3], AND NATALIA LYSENKO[4]

[1] Stony Brook University
[2] University of Illinois Urbana-Champaign
[3] Arizona State University
[4] Orel State University

ABSTRACT

This study compares Russian native speakers and adult L2 learners on their ability to locate new information and contrastive foci in SVO and OVS sentences during silent reading and listening comprehension. Participants completed a reading task and a listening task identifying the focused word using context and prosodic cues. L1 English L2 Russian speakers performed accurately in reading, but in listening half of the items contained a mismatch between focus and nuclear pitch prominence. Participants consistently identified prosodically accented words in contrastive focus, but showed greater variability with new information focus in sentence-final position. This may reflect differences between pitch accents marking contrastive vs. new information foci, compounded by the prosody–focus mismatch in half of the auditory stimuli.

KEYWORDS second language acquisition · focus · prominence · cross-linguistic influence

1 INTRODUCTION

Information status is traditionally linked to the relative familiarity with, and/or accessibility of, a word referent in its contextual environment, or discourse (Arnold et al. 2013). If one regards referent accessibility as a continuous measure, the two end points of the accessibility continuum would be a word introducing a brand-new referent, which the reader or the listener perceives as new information, and a word referring to some highly accessible information based on (repeated) prior mentions. This means that accessible information is information already known to the reader/listener, i.e. given. Information structure (IS) models motivate a systematic distinction between the information which is given vs. new (for an example of a non-binary model of information accessibility, see Baumann & Riester 2013), but also include such categories as focus and contrastiveness (Pierrehumbert & Hirschberg 1990).

The given information connects the utterance with the preceding discourse, while the new information updates the cognitive state of the listener or the reader, advancing their knowledge further. As a result, some languages opt for ordering these information categories such that given information linearly precedes novel information, giving rise to utterance-level ordering preferences driven by considerations of information status. Relevant experimental evidence comes from an ERP study by Pyykkönen et al. (2011) that found that in German ditransitive verb constructions, listeners expect to have discourse-given entities mentioned prior to the discourse-new entities. In a similar vein, Kaiser & Trueswell (2004) found that in Finnish, a language which allows for various word order permutations in addition to the basic SVO order, starting a sentence with an object is only felicitous when the latter is associated with given or established information,

while the subject of the sentence is previously unmentioned. The findings of Pyykkönen et al. (2011) and Kaiser & Trueswell (2004) corroborate that considerations of relative information accessibility may have repercussions for constituent linearization, such that placing a word in a position that accords with its information status may optimize the flow of information and facilitate comprehension.

While changes to the basic order of phrasal constituents can be afforded, primarily, by the so-called free word order languages (e.g., Finnish and Russian) but remain largely unavailable in rigid constituent order languages (e.g., English), word-level prosodic expression presents another domain that in many languages is sensitive to information structural distinctions. In fixed word order languages, acoustic-prosodic expression presents an important way of communicating the information status of a word (e.g., Cole et al. 2019). Prosodic expression of IS categories "new" and "given" usually concerns perceptible changes in the voice quality of the speaker and may involve segmental duration and intensity, as well as the use of supra-segmental cues, such as changes in the fundamental frequency, or pitch. Typically, discourse-new words are associated with a perceptual rise in pitch, known as a pitch accent, often accompanied by greater duration of the stressed vowel in the accented word (Breen et al. 2010). In a survey of more than twenty acoustic properties of the speech signal that can potentially correlate with different categories of information structure in English, Breen et al. (2010) found that greater intensity, longer duration, and higher mean and maximum f0 are all reliable predictors of the location of new information focus. Contrastive focus presents a distinct category of focal information in discourse (English: Katz & Selkirk 2011; Russian: Bryzgunova 1980), whereby the referent of the focused word is singled out from a set of contextually available alternatives (Rooth 1992). Of relevance to the present study is that in languages which mark focus by means of prosody, contrastive focus is prosodically differentiated from new information focus, which renders it both perceptually distinct and, possibly, more perceptually prominent (see Bishop 2012 on English and §2 below for a discussion of contrastive focus prosody in Russian). Whether the prosodic correlates of contrastive focus translate into a more accurate focus identification during listening remains an open question (Breen et al. 2010).

Prosodic expression of information status is widely attested in many (but not all) languages which demonstrate flexible word orders: f0 and duration are described as key acoustic-prosodic correlates of information structural distinctions in European Portuguese, Greek, Dutch, Italian, and Romanian (Frota 2002, Swerts 2007, Swerts et al. 2002). The present study examines L1 and L2 identification of discourse-new and contrastive information in Russian, a relatively free word order language, which also uses acoustic-prosodic means to express word-level IS cues. In what follows, we will use the terms "new information focus" when referring to the IS category "discourse-new" and "contrastive focus" when referring to contrastive information. We will proceed by reviewing the linguistic means which may be used to express sentence focus in Russian (§2) and briefly discuss cross-linguistic experimental evidence pertaining to L2 acquisition of focus identification (§3 and §4). We will then turn to the present study (§5 and §6), discuss the implications of our findings (§7) and outline some possible directions for future work (§8).

2 REPRESENTING IS DISTINCTIONS IN RUSSIAN

Russian is a free word order language in which about 20% of spoken utterances deviate from the canonical SVO order[1] (Lobanova 2011, Sekerina 2003). In Russian, prosodic effects related to the information status of a discourse referent are evident in its patterns of pitch-accenting. This entails the accentuation of new, focused information and the deaccenting of given information (Jasinskaja 2014, Neeleman & Titov 2009). By default,

---

[1]We use the term "canonical word order" to refer to the default constituent order in an all-new information sentence.

the Nuclear Stress Rule (Chomsky & Halle 1968) aligns the main phrasal prominence, known as the nuclear pitch accent, with the rightmost prosodic phrase boundary (Kondrashova 1997). Russian is similar to other pitch accenting languages in that rendering a non-sentence-final word focused leads the speaker to align the nuclear pitch prominence with the focused constituent non-sentence-finally. Russian may optionally deploy changes in word order to express information status and relative information prominence (Luchkina & Cole 2016, 2021).

To illustrate, in (1), the answer to the subject question 'Who scared the wolf?' can either use non-canonical OVS order, shown in (1-b), where the focused subject noun *lisa* 'fox' appears sentence-finally and thereby aligns with the sentence-final nuclear pitch accent, or else canonical SVO order, with a shift of the nuclear pitch accent to the preverbal position (1-c). Similarly, the object question 'Whom did the fox scare?', in (2), can be answered either with SVO order and sentence-final nuclear pitch accent (2-b), or with non-canonical word order and shifted nuclear pitch accent (2-c).

(1)   a.   Kto   napugal volka?   (Russian)
             who.NOM scared   wolf.ACC
             'Who scared the wolf?'
      b.   Volka   napugala LISA.
             wolf.ACC scared   fox.NOM
      c.   LISA   napugala volka.
             fox.NOM scared   wolf.ACC
             'A fox scared the wolf.'

(2)   a.   Kogo   napugala lisa?
             who.ACC scared   fox.NOM
             'Whom did the fox scare?'
      b.   Lisa   napugala VOLKA.
             fox.NOM scared   wolf.ACC
      c.   VOLKA   napugala lisa.
             wolf.ACC scared   fox.NOM
             'The fox scared a wolf.'

Russian further supports differential prosodic marking of distinct types of focal information, including non-contrastive, new information foci and contrastive foci. An investigation by Bryzgunova (1980) revealed that each focus type is associated with a distinctive prosodic realization in speech, based, primarily, on the shape of the f0 contour over the accent bearing unit. Specifically, new information focus receives an HL* accent distinguished by a prominent falling contour over the stressed syllable, whereas the contrastive focus receives an H*L accent and bears a more pronounced rise and a higher pitch peak over the stressed syllable with a subsequent convex-shaped fall (see Jasinskaja 2014 for more discussion).

Differential accenting patterns in relation to contrastive and non-contrastive foci are not unique to Russian, as pitch accenting languages differ with respect to the contour shapes associated with specific focus types (e.g., L+H* marks contrastive focus in English; see Bishop 2012 for more discussion). Native listeners, then, must be equally sensitive to both constituent linearization and prosodic patterns of expressing focal information in Russian. This was previously demonstrated by Luchkina & Cole (2021), who conducted a perceptual rating task in which a group of adult native listeners were instructed to identify words which they perceived as prominent. As expected, participants in this study were more likely to rate prosodically augmented words, as well as the words occurring in ex-situ sentence positions, as prominent. Both these properties help draw a perceptual distinction between discourse-new or otherwise prominent information (e.g., due to contrastive focus or emphasis) on the one hand and some less salient information categories, such as information which presents common knowledge or is readily inferable

from context, on the other.

3 KNOWN L2 ACQUISITION CHALLENGES

While word order and prosodic means appear reliable cues to IS distinctions in native comprehension, they present a known challenge in adult and heritage language acquisition, possibly due to the greater complexity of grammatical structures at the interface between two or more linguistic domains, e.g. Luchkina et al. (2024). The Interface Hypothesis originally developed in Sorace & Filiaci (2006) is one theoretical account which emphasizes the acquisition challenge associated with interface phenomena involving syntax in combination with other language domains, such that the target interface structures may not be completely acquired by even advanced-level learners. This is predicted to be in stark contrast with the target structures that are purely syntactic in nature and are therefore fully acquirable. Presenting an area of maximal difficulty, the interface phenomena (e.g., anaphor resolution and focus identification) are associated with optionality in relation to the target grammar, as well as "residual indeterminacy" (Sorace & Filiaci 2006:339). Of interest to the present study, Sorace (2011:5) discusses the so-called grammar-external interfaces, including the syntax-pragmatics interface, as challenging not only for adult learners but also for bilingual L1 acquisition, due to the involvement of "contextual information external to the grammar".

While the IH predictions have not been regarded as entirely unproblematic, due to patterns of individual variability among L2 learners, and due to a lack of clarity regarding the nature of the non-target-like performance (see White 2011 for a helpful discussion of representational vs. processing difficulties in advanced level learners), evidence consistent with its main tenet, namely, that external interfaces resist acquisition and are vulnerable even during the early stages of language attrition, continues to accrue.

Ionin et al. (2023) tested participants in a bimodal task with dialogues shown in (1) and (2) above, which consisted of either a subject question (1-a) or an object question (2-a), and an answer that always had nuclear pitch accent sentence-finally (1-b), (2-b). For subject questions (1-a), native speakers (NSs) of Russian rated (1-b) high and (2-b) low. For object questions (2-a), NSs rated (2-b) high and (1-b) low. The performance by the native speaker group revealed that listeners were sensitive to the location as well as the prosodic expression of the focused word. It was when these factors clashed with the IS that the NSs correctly identified the answer sentences as infelicitous. Heritage speakers of Russian performed similarly to the native monolinguals. In contrast, English-speaking adult L2 learners of Russian rated the canonical SVO order (2-b) above OVS order (1-b) in answer to both (1-a) and (2-a), despite the fact that sentence-final stress on the object in (2-b) clashed with focus on the subject in (1).

4 THE LOCI OF NON-TARGET-LIKE PERFORMANCE

The question that arises from the findings reported in Ionin et al. (2023) relates to the source of L2ers' non-target performance. One possibility is that the locus of the difficulty is perception-based, i.e. due to (i) learners' difficulty locating the focused word using prosodic cues; alternatively, the difficulty may be due to (ii) learners' difficulty integrating the word order and prosodic cues with the propositional content of the utterance in order to achieve the intended reading, as would be predicted by the IH. Both these possibilities are further explored in the present study, using the experimental stimuli originally developed for Ionin et al. (2023) in new perceived prominence rating tasks and with new participant groups.

The perception-based account is compatible with the findings reported in studies which investigated cross-linguistic differences in focus marking strategies (see Cole et al. 2019 for a discussion of correlates of phrasal prominence and its relationship with IS in Spanish, English, and French), as well as acquisition studies which examined perception

of focus/phrasal prominence by adult L2 learners of typologically different languages. To illustrate, in a perception-production investigation by Ortega-Llebaria & Colantoni (2014), the authors found that the focus marking strategy instantiated in the learners' L1 may give rise to long lasting negative transfer effects observed in even highly proficient L2 learners. One experiment reported in that paper tested perception of English sentence focus by L1 Spanish and L1 Mandarin L2 English learners. Both groups were instructed to determine the location of the contrastively focused word in sentences presented with or without context. The study found that the L1 Mandarin participants were on target, and virtually indistinguishable from the English NSs despite being less proficient than the L1 Spanish group; this was likely due to positive transfer from their L1, which also uses prosodic prominence to express focus. The L1 Spanish group, despite being more proficient overall, was significantly less accurate, possibly because Spanish uses word order to align the focused constituent with the sentence-final position where it receives the nuclear pitch accent; as expected, L1 Spanish speakers' accuracy was greater when the focused constituent was the post-verbal object (no stress shift required) than when it was the pre-verbal subject (stress shift required), again pointing to robust L1 transfer effects due to differences in focus marking strategies used in the L1 and the L2.

Furthermore, learners' proficiency can influence their ability to accurately perceive IS-based distinctions in the TL. Lee et al. (2019) investigated perception of variably positioned phrasal prominence which served to express contrastive focus by L2 English learners whose L1s were Korean and Spanish. Results revealed greater accuracy in the performance of L1 Korean learners of English, attributed to the similarities in the use of prosodic cues marking phrasal prominence in Korean and English. Both L2 groups demonstrated effects of TL proficiency which positively correlated with accuracy irrespective of the participant's L1.

To summarize, adult L2 learners' ability to identify the focused word may be compromised by mismatches in the use of focus marking cues in the L1 and in the L2, as well as the learner's proficiency level. Under the *perception-based account*, factors such as these may have contributed to the non-target-like comprehension and over-acceptance of non-felicitous focus readings reported in Ionin et al. (2023) for L2 Russian. Under the *IH account*, however, the locus of L2 learners' non-target-like performance is not perception-based, but, rather, is due the difficulty in acquiring the use of word order and prosodic cues in the TL to achieve desired IS effects, such as focus. We proceed by critically evaluating each of these accounts using L2 Russian as the test case.

5 METHOD

5.1 DESIGN

The present study (a) compares Russian NSs and adult L2 learners on their ability to locate the focused word in sentences like (1-b) and (2-b) during silent reading, i.e. in the absence of acoustic-prosodic cues; and (b) gauges the relative sensitivity to prosodic cues which signal the location of the main phrasal prominence in these populations. To achieve these goals, we conducted a silent reading task and a listening task. For both tasks, we used stimuli which were created by crossing two levels of the factor "word order", SVO and OVS, with two levels of the factor "focused noun", subject noun and object noun, to create new information focus target items. Both tasks presented dialogues in which either a subject question (3-a) or an object question (4-a) was followed by either an OVS answer (3-b) or an SVO answer (4-b). Subject questions placed the subject in new information focus, while object questions placed the object in new information focus. The nuclear pitch accent was invariably on the sentence-final word (the subject in OVS order (3-b) and the object in SVO order (4-b)), creating an equal number of prosodically felicitous dialogues, (3-a)–(3-b) and (4-a)–(4-b), and prosodically infelicitous ones, (3-a)–(4-b) and (4-a)–(3-b). We used a total of 96 items, 24 for each type of dialogue.

(3)   a.   Kto   napugal volka?   (Russian)
                who.NOM scared   wolf.ACC
                'Who scared the wolf?'
         b.   Volka   napugala lisa.
                wolf.ACC scared   fox.NOM
                'A fox scared the wolf.'

(4)   a.   Kogo   napugala lisa?
                who.ACC scared   fox.NOM
                'Whom did the fox scare?'
         b.   Lisa   napugala volka.
                fox.NOM scared   wolf.ACC
                'The fox scared a wolf.'

In order to test the perception of contrastive focus, known to be intonationally distinct from non-contrastive new information focus in Russian (see discussion in §2 above and Bryzgunova 1980), we used 96 contrastive focus items, each consisting of question-answer dialogue; a sample question is given in (5-a), with four different answer types in (5-b)–(5-e). In the contrastive focus items, any word in the answer sentence could be nuclear pitch accented, yielding a felicitous or an infelicitous reading presented in the listening task: for example, (5-b) and (5-c) are felicitous responses to (5-a), while (5-d) and (5-e) are infelicitous (the accented word in each answer sentence appears in SMALL CAPS). These items also featured more varied constituent orders, e.g. in (5-b) vs. (5-c), prepositional phrases, and modifiers. Cumulatively, the design features of the contrastive focus items obviated the need for fillers, for either task.

(5)   a.   Vy včera   hodili   v kino?   (Russian)
                2PL yesterday go.PST.PL to cinema
                'Did you got to the cinema yesterday?'
         b.   Net, my hodili   v TEATR.
                no   1PL go.PST.PL to theatre
         c.   Net, my v TEATR hodili.
                no   1PL to theatre go.PST.PL
         d.   Net, MY v teatr   hodili.
                no   1PL to theatre go.PST.PL
         e.   Net, my HODILI   v teatr.
                no   1PL go.PST.PL to theatre
                'No, we went to the theatre.'

All of the test items were taken from the original acceptability judgement task developed by Ionin et al. (2023). Use of the same stimuli across studies allows us to directly investigate the reasons for L2 speakers' poor performance in Ionin et al. (2023). The auditory versions of the items had been recorded by two female native Russian speakers. Only felicitous dialogues were recorded in order to avoid the production of unnatural intonation in infelicitous dialogues; all questions and answers were then separated and spliced together to form the infelicitous dialogues.

Testing was administered via Qualtrics. Four experimental lists were built, each of which consisted of 24 new information items and 24 contrastive focus items. Token sets were distributed across the four lists using the Latin-square design. The dialogues in each list were presented in a pseudo-randomized order. All participants completed the silent reading task first, followed by the listening task. Both tasks opened with three practice trials but followed a different item list. In addition to the two main tasks, L2ers completed a cloze test as a measure of their Russian proficiency (not reported in the present paper) and a case check which helped us monitor if they used case marking,

rather than word order, in determining the subject and the object in the test sentences.[2] The case check task consisted of ten written three-member transitive sentences, each followed by a comprehension question of the type 'Who did X'. Half of the sentences were subject-initial (SVO), such that the correct answer was the sentence-initial subject. The remaining sentences were all object-initial (OVS), such that the correct answer to the comprehension question was the sentence-final subject. The case check sentences were constructed such that each comprehension question could be plausibly answered using either the subject or the object noun and each object noun was unambiguously case-marked.

In the **silent reading task**, participants were presented with 48 short dialogues of the form described above. The question sentence was presented auditorily and the answer sentence was presented in writing. Participants were instructed to select the most prominent word in the answer sentence, by clicking on the desired word on their test screen. In the absence of any prosodic cues in the answer sentence, only contextual cues could be used to determine which word is the most prominent: after a subject question (3-a), the subject noun ('fox') in the answer is prominent, both in OVS order (3-b) and in SVO order (4-b); conversely, after an object question (4-a), the object noun ('wolf') in the answer is prominent, again both in OVS and SVO orders.

In the **listening task**, each dialogue was presented auditorily. Each test item consisted of an audio recording of one complete dialogue followed by the answer sentence from that dialogue presented in writing on the same screen. Each content word in the answer sentence was clickable. Participants were instructed to pay close attention to the prosody of the answer sentence, and to select the word which stood out the most based on word-level prosodic cues. Participants were informed that in some dialogues, the choice of the prosodically prominent word could seem odd and that in such cases, they were to pay close attention to the intonation and disregard contextual appropriateness. Participants selected their answers by clicking on the desired word on their test screen.

5.2 PARTICIPANTS

The NSs group consisted of thirty-five participants (mean age = 37.1; 10 males). All NSs resided in Russia at the time of participation and had acquired Russian from birth. Almost all participants reported being functionally monolingual with the exception of one speaker who identified as a Russian–Ukrainian bilingual. All participants reported Russian as their native language.

The L2 sample included 28 adult English-Russian bilinguals (mean age = 20.6; 9 males). The average age of exposure to English was 2.0 years of age; the average age of exposure to Russian was 13.9 years of age. All participants resided in the US at the time of testing and reported English as their primary language of communication. Eight participants self-identified as heritage Russian speakers due to (some) exposure to Russian via one or both parents. Twenty-two participants were enrolled in college-level courses of Russian as a foreign or heritage language at US colleges at the time of testing. Because the heritage learners all declared English to be their stronger (dominant) language and because their performance in the experimental tasks was overall rather similar to that of adult L2ers, we refer to all 28 participants as the "L2 group" in the remainder of this paper.

5.3 PREDICTIONS

The present section lays out testable predictions for the silent reading and the listening task first for NS participants and then for L2ers. We predicted that (a) in the silent reading task, NSs would select the focused noun as prominent regardless of its position

---

[2]See Ionin et al. (2023) for more information on these proficiency measurements, and Luchkina et al. (2021) for the specifics of the cloze test.

in the sentence. If prediction (a) is borne out, in answer to subject questions, such as (3-a), the subject should be identified as prominent (LISA 'fox' in (3-b) and in (4-b)), and in answer to object questions, such as (4-a), the object (VOLKA 'wolf' in (3-b) and in (4-b)). The contrasted word (TEATR 'theatre' in (5-b) through (5-e)) should be identified in the contrastive focus items.

Turning to the listening task, recall that the accented noun was always sentence-final in the new information focus items; the location of the accented noun in the contrastive focus items was variable and required that listeners pay careful attention to the acoustic-prosodic expression. Finally, in half of all dialogues, the accented word and the word in focus were different, further motivating the listeners to pay close attention. We predicted (b) that NSs would accurately identify the accented word as prominent regardless of its information status in the listening task. If prediction (b) is borne out, test dialogue felicity should have no significant effect on the rate of prominent word identification among the NS listeners.

Our predictions for the L2 participants were more tentative due to scarcity of relevant prior work. Based on the results reported by Ionin et al. (2023), we expected (c) to observe greater accuracy in the performance of the L2ers who could reliably distinguish between SVO and OVS constituent orders by appropriately interpreting the case suffix in subject and object nouns in the case-match task; if (c) is borne out, the L2ers' performance in the case-match task should be predictive of their performance in the silent reading as well as the listening tasks; (c) was additionally tested by gauging the effect of the L2 participants' age of acquisition of Russian (henceforth, AoA) on their accuracy rate in each completed task. Furthermore, given the L2ers' preference for canonical order reported in Ionin et al. (2023), (d) we predicted that the success rate on items testing new information focus would be greater when the answer follows canonical SVO order. Finally, we tentatively predicted (e) target-like performance in the silent reading task (the prominent element is the one that answers the question), as well as good accuracy in the listening task, due to positive transfer from the L1/dominant language (English), which uses prosodic accenting to express focus. If prediction (e) is borne out, test item felicity should have no significant effect on the rate of prominent word identification among the L2 listeners.

6 RESULTS

The NS and L2 participants exhibited similar patterns in the silent reading and listening tasks. We therefore report the accuracy means and results of the inferential analyses by item type (contrastive focus first and new information focus next) for the two participant groups.

All participants demonstrated good accuracy when asked to identify the prominent word in contrastive focus items, independent of the task modality. Recall that in the Silent Reading Task, context served as the primary (and only) cue to focus. Native speakers demonstrated at ceiling, 99% accuracy and were unanimous in their identification of the contrastively focused word as the most prominent word in the answer sentence. Impressively, the L2 participants performed at 95% accuracy revealing that they, too, understood the dialogues and, just like the NSs, perceived the contrastively focused word as prominent.

While *all* the dialogues presented in the Silent reading task were felicitous (since no prosodic cues were provided in the answer sentence), this was not the case in the Listening task, where both the question sentence and the answer sentence were presented auditorily. In half of the contrastive focus dialogues, the choice of the prosodically accented word in the answer sentence clashed with the context set up in the question sentence. The listening task specifically asked that the respondent selects the word which the model speaker highlighted by means of prosody, irrespective of the felicity effect. Despite the instructions, the accuracy rates obtained in the listening task reveal that overriding the felicity effect was not entirely unproblematic for both participant groups.
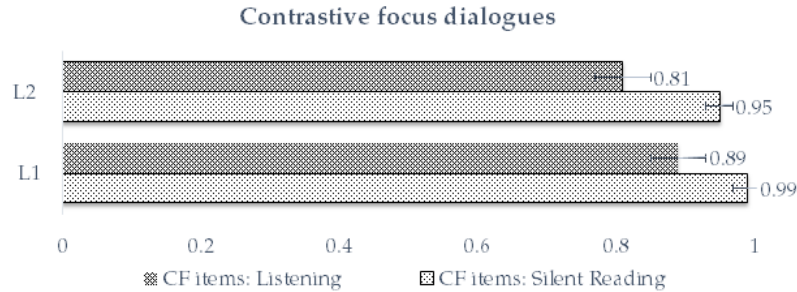
Contrastive focus dialogues

Figure 1: Mean accuracy on contrastive focus items in the Listening task and the Silent Reading task. Error bars display standard error. L1 = native speaker group; L2 = L2 group

Figure 1 summarizes the accuracy means obtained for contrastive focus items (felicitous and infelicitous shown together). The NSs' mean accuracy reached 96% on the felicitous items and 83% on the infelicitous items. The L2ers were on average 94% accurate on the felicitous items but only 68% on the infelicitous items.
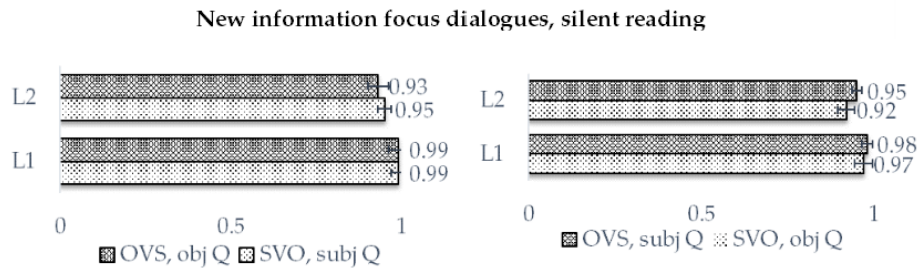
New information focus dialogues, silent reading

Figure 2: Mean accuracy on new information focus items in the Silent Reading task. Error bars display standard error. L1 = native speaker group; L2 = L2 group

Results obtained for the new information focus items differed by task modality. All participants achieved excellent accuracy in the Silent Reading task, as shown in Figure 2. The accuracy means obtained for new information focus items in the listening task are summarized in Figure 3. On-target performance in this task required that the sentence-final, pitch accented noun be selected regardless of the felicity effect. For felicitous dialogues, NSs identified the sentence-final word as prominent 56% of the time, but for infelicitous - only 33% of the time. L2ers selected the sentence-final word as prominent in 59% of felicitous dialogues but just in 22% of infelicitous dialogues.

The inferential analyses were carried out separately for contrastive focus and new information focus items due to the many differences in item design. All mixed effects regression analyses reported below included random effects (intercepts) for Participant and Item.

We implemented four mixed effects logistic regression models using responses to contrastive focus items. Model A was run on the silent reading task data obtained from all participants and included the fixed effect "participant status" (native vs. L2). This analysis returned a main effect of participant status ($z = 5.21$, $p < .001$), indicative of NSs being overall more accurate. Model B presented a follow up analysis implemented using data from L2 learners only, which allowed us to incorporate two additional main effects, the Russian AoA, and the case check score. Model B revealed a significant main

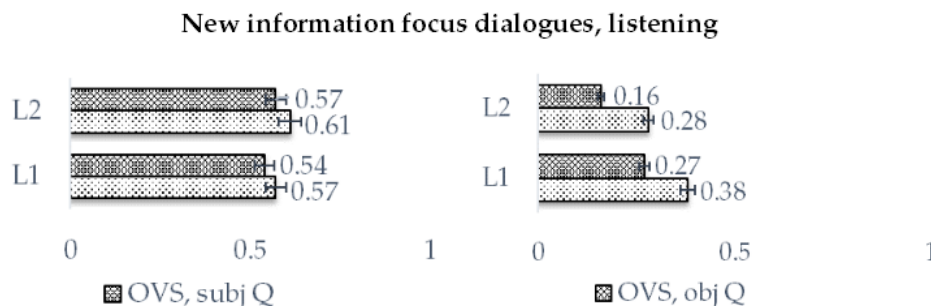New information focus dialogues, listening

Figure 3: Mean accuracy on new information focus items in the Listening task. Infelicitous dialogue types are presented on the right. Error bars display standard error. L1 = native speaker group; L2 = L2 group

effect of the Russian AoA, which was negatively associated with the log likelihood of a correct response ($z = -3.11$, $p = .002$). Models C and D were run on the listening task data and included the fixed effects "token" (4 levels, corresponding to the four different answer options, see (5-b)–(5-e)) and "participant status". Model C failed to detect a main effect of participant status but returned a significant main effect of "token". For all tested participants, the log likelihood of a correct response was significantly lower for infelicitous items (two out of four tokens in each set: $z = -3.76$, $p < .001$; $z = -2.49$, $p = .01$). Model D was fit using L2 data only, and additionally revealed significant main effects of the Russian AoA ($z = -2.30$, $p = .02$) and the case check score ($z = 2.35$, $p = .02$). Both effects point to the comparative advantage of the learners with earlier AoAs to Russian, as well as better performance on the case check task.

Now, turning to the performance on new information focus items, because the dialogues in this category systematically crossed question type and word order in the answer sentence, both these factors were included as fixed effects and were also interacted with each other. Model E was implemented using the silent reading task data from all participants. Participant status (native vs. L2) was the only significant main effect in this analysis ($z = 4.54$, $p < .001$), indicative of more accurate performance by the NSs. Model F was fit using L2 data only and returned a significant main effect of the case check score ($z = 3.37$, $p = .001$). Models G and H were implemented using data from the listening task. Model G failed to detect a significant main effect of participant status but returned a significant interaction between word order and question type. Specifically, the log likelihood of an accurate response was lowest for dialogues which consisted of an object question followed by an OVS answer, compared to subject questions followed by OVS answers ($z = 6.1$, $p < .001$), and object questions followed by SVO answers ($z = 7.27$, $p < .001$). Comparing the two infelicitous dialogue categories, the log likelihood of a correct response for subject questions followed by SVO answers was still greater than that for object questions followed by OVS answers ($z = 2.34$, $p = .02$). In addition to the significant interaction between question type and word order in the answer sentence (same pattern as in model G), model H returned significant main effects of the Russian AoA ($z = -4.15$, $p < .001$) and the case check score ($z = 4.17$, $p < .001$).

7 DISCUSSION

The present study seeks to test two proposals which may account for difficulties demonstrated by L1 English L2 learners of Russian when judging the relative acceptability of word order-prosody configurations used in the expression of information structural distinctions in the TL (Ionin et al. 2023). According to one of these proposals, the locus

of the L2 learners' non-target-like performance may be perception-based, related to difficulties detecting or decoding the relevant acoustic-prosodic cues used to mark sentence focus in Russian. Under a different account, the acquisition challenge transcends beyond the matters of perception and emerges when the L2 interfaces a purely syntactic mechanism (word order) with an external domain (Information Structure), to yield a desired interpretation. To probe into each of these possibilities, we tested (i) perception of distinct word order – prosody configurations deployed as ways of expressing new information focus, as well as (ii) perception of prosodic cues marking contrastive focus, in monolingual Russian speakers and in English-dominant adult learners of Russian as a second or heritage language. Additionally, we manipulated whether prosodic accenting used to mark sentence focus was contextually supported, in order to encourage participants to pay even closer attention to the acoustic-prosodic expression.

For the **silent reading** task, we predicted that NSs would select the focused noun as prominent regardless of its position in the sentence, whereas the mean accuracy of the L2ers would hinge on their ability to reliably distinguish between the SVO and the OVS orders, tested in the case-match task. We found that when no acoustic-prosodic cues were provided, words placed in contrastive focus or in new information focus were reliably perceived as prominent by both the NS group and the L2 group. Very high accuracy rates obtained from all participants indicated that they were able to pick out focused words based on contextual cues alone. While the L2ers' performance on the case check score was not predictive of their overall accuracy in the silent reading task (relatedly, the constituent order in the target sentence did not play a role), their AoA to Russian was indicative of greater rate of success on focus identification for those with an early age of exposure to the Russian language.

In the **listening** task, participants had to pay careful attention to the use of prosodic cues in the model speaker's read production and to select the most prosodically prominent word. We predicted that NSs would accurately identify the accented word as prominent regardless of its information status and that the test dialogue felicity should have no effect on the rate of prominent word identification. For L2ers, we predicted to observe good accuracy in the listening task, due to positive transfer from the L1/dominant language (English), which also uses prosodic accenting to express focus. We additionally predicted greater focus identification accuracy on SVO stimuli, as well as for participants with higher scores in the case check task.

Both groups were successful when asked to detect the location of **contrastive focus** based on prosodic cues; the success rates were particularly high when a contrastive focus reading was supported by means of context. The accuracy rates decreased by about 10% on infelicitous items, suggestive of emerging context effects which lead some participants to select the prominent word based on context, rather than on acoustic-prosodic cues.

A rather different pattern of results was obtained in the listening task on test items featuring a noun placed in **new information focus**. We systematically varied the word order in the target (answer) sentences and combined each such sentence with two question types, subject and object. The location of the nuclear pitch accent remained sentence-final in all items. Our new information focus item design yielded some well-formed dialogues in which the focused noun was sentence final and therefore aligned with the nuclear pitch accent. The other half of the dialogues, however, was ill-formed in that the focused noun was sentence-initial and thereby failed to align with the main sentence prominence landing site. The obtained accuracy rates were again similar for both participant groups but were otherwise in stark contrast with the previously mentioned results. Specifically, the accuracy was rather low (at 56–59%) for the well-formed new information focus items and decreased further (22–33%) for the items in which IS clashed with prosody. Thus, in the listening task, both groups were more accurate on felicitous dialogues featuring the focused noun in sentence-final position than on infelicitous ones in which the nuclear pitch was sentence-final, but the focused noun was sentence-initial. This finding was not what we had predicted but can be justified given that in head-initial languages, including

English and Russian, the most embedded constituent in a clause (typically, the object) is interpreted as focused and therefore receives the nuclear pitch accent (English: Selkirk 1995; Russian: Kondrashova 1997).

An even less anticipated result is that both groups demonstrated different accuracy rates obtained for contrastive foci (greater accuracy) and new information foci (lower accuracy) in the listening task. We hypothesize that this may be due to participants' task-induced expectation to observe variable as well as more pronounced accenting patterns throughout the test. Following Bryzgunova (1980), Jasinskaja (2014) discusses a number of qualitative differences in the quality of the pitch accents associated with new information focus (labeled HL*) and "non-neutral", contrastive focus (labeled H*L) in Russian. A critical difference between the two accent types is that the former (observed in canonical SVO sentences with sentence-final nuclear pitch accent) is compatible with two readings, including broad focus (all new information) and narrow focus on the sentence-final word. Prosodic accenting coincidental with contrastive focus, on the other hand, is unambiguously aligned with one unique sentence constituent. Furthermore, according to Bryzgunova (1980), HL* is realized on the sentence-final word and therefore has a relatively low pitch excursion. Additionally, the HL* contour may be preceded by one or more pre-nuclear pitch accents occurring earlier in the sentence, further dampening its perceptual salience. The H*L contour, on the contrary, may align with any (non-final) word and appear more salient due to its high f0 peak; it may also be preceded by fewer pre-nuclear pitch accents, which are optional in sentences featuring contrastive focus. Cumulatively, these properties account for a greater pitch excursion in H*L compared to HL*, pointing to possibly dissimilar perceptual saliency of the two accent types (see Wagner et al. 2010 and Bishop 2012 for similar arguments on English). Such differences in perceptual saliency of the tested accenting patterns, then, could have given rise to indeterminacy in prominence ratings during the listening task. Additionally, the presence of context cues which half of the time clashed with the focus reading of the prosodically prominent word could further destabilize listeners' judgements. These assumptions call for further investigation, as briefly outlined in §8.

Prior research has documented effects of TL proficiency on perception of various focus types, including new information focus and contrastive focus (see §3 for details). In the present study, native status yielded a small advantage in terms of the overall performance accuracy, which further led us to examine the effects of L2 participants' age of acquisition of the TL and their performance on an independent case check task designed to gauge the ability to correctly use case marking, rather than word order, in the TL to determine the grammatical function of a noun (subject or object). Inferential analyses revealed that L2ers' AoAs and case check results mattered the most in the listening modality, suggesting that listeners' responses reflected their interpretation of the target sentences rather than a mere reliance on the prosodic cues. These results tentatively point to a greater difficulty of the listening task, requiring that the L2 participants jointly attended to context cues, word order, and prosody.

Nevertheless, the two participant groups performed very similarly. The results of the listening task show that the L2 participants were highly accurate at locating contrastive focus. These results clearly contradict the perceptual account, according to which difficulties in the acquisition of interface phenomena are perception-based. Turning now to our findings concerning new information focus, results of the silent reading task revealed that our L2 participants recognized the relationship between word order and information structure. In the listening task, both groups had an easier time identifying the prosodically prominent word when it was also the word bearing focus. This result of the listening task suggests that the L2ers were able to perceive prosodic prominence, as well to integrate it with the contextual cues (the preceding question), albeit with lower accuracy.

How can the target-like performance of the Russian learners tested in the present study be reconciled with the findings reported by Ionin et al. (2023)? In the auditory

acceptability judgement task conducted by Ionin et al. (2023), the proficient Russian L2ers showed stable preference for the canonical SVO order even when the sentence-final stress on the object clashed with focus on the subject. The results reported by Ionin et al. (2023) are compatible with the IH account, whereas the present study finds rather similar performance patterns in the L1 and the L2 Russian participants. One critical difference between the Ionin et al. (2023) study and the present one is that the former taps into metalinguistic knowledge by asking listeners to rate the relative acceptability of the WO-prosody configurations corresponding to well-formed and ill-formed focus readings, whereas the present study asks participants to pick out the prosodically prominent word regardless of the well-formedness considerations. Our results, therefore, test focus identification and tap into perceptual sensitivity more so than those of Ionin et al. (2023). It is also possible that an acceptability judgement task is more computationally taxing than a perceived prominence identification task and therefore presents a better test case for the IH hypothesis, which appeals to computational load as one of the loci of L2 learners' non-target-like performance. Our own results remain rather inconclusive with regard to the IH. On the one hand, results of the silent reading task reveal that the L2ers were able to compute IS relations despite the syntactic manipulations in the stimuli permissible in the TL but not in the L1. On the other hand, in the listening modality, responses may have been influenced by the infelicity of some of our items, and also by the fact that our tasks included instances of contrastive focus and new information focus and therefore exposed participants to patterns of prosodic expression differing in terms of their perceptual salience and underlying prosodic correlates.

8 CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH

Adult L1 English L2 Russian speakers demonstrated on-target performance when asked to locate a focused word based on context cues and in the absence of acoustic-prosodic information. This suggests that they correctly understood the use of SVO and OVS constituent orders leading to on-target interpretation of IS status. When prosodic cues were provided, participants readily identified the nuclear pitch accented word placed in contrastive focus. Participants displayed considerably more variability in the perception of prosodically prominent words in the sentence-final, nuclear-pitch accented position. This was possibly due to qualitative differences in the nature of pitch accents associated with contrastive foci and new information foci, compounded by the fact that half of the auditory stimuli in the listening task clashed prosody and focus.

The present study sets the stage for further work investigating the acquisition of the word-order prosody interface in the L2 acquisition of Russian foci. Taking into account the method used in the present study, a logical next step is to conduct a listening task with target SVO and OVS sentences presented without context. This would allow for a more straightforward test of listeners' sensitivity to prosodic cues marking new information focus in Russian. A closer look (i) at differences in the prosodic expression of new information focus vs. contrastive focus and (ii) at how the prosodic cues used to manifest these foci types compare in the L1 (English) and the TL is also in order, to better account for individual variability in the perception of prosodic cues found with respect to new information focus in the present study. Lastly, future work should aim to more systematically examine perception data from L2ers of differing proficiency levels, in order to see whether sensitivity to contextual cues is acquired before, or hand-in-hand with, sensitivity to prosodic cues.

CONTACT

Tatiana Luchkina — *<tatiana.luchkina@stonybrook.edu>*
Tania Ionin — *<tionin@illinois.edu>*

Maria Goldshtein — *<mgoldsht@asu.edu>*
Natalia Lysenko — *<natalia.lysenko@inbox.com>*

ABBREVIATIONS

1	first person	NOM	nominative
2	second person	PL	plural
ACC	accusative	PST	past
DAT	dative		

REFERENCES

Arnold, Jennifer E., Elsi Kaiser, Jason M. Kahn & Lucy Kyoungsook Kim. 2013. Information structure: Linguistic, cognitive, and processing approaches. *Wiley interdisciplinary reviews: Cognitive science* 4(4). 403–413.

Baumann, Stefan & Arndt Riester. 2013. Coreference, lexical givenness and prosody in German. *Lingua* 136. 16–37.

Bishop, Jason. 2012. Information structural expectations in the perception of prosodic prominence. In Gorka Elordieta & Pilar Prieto (eds.), *Prosody and meaning*, 239–270. Berlin: Mouton de Gruyter.

Breen, Mara, Evelina Fedorenko, Michael Wagner & Edward Gibson. 2010. Acoustic correlates of information structure. *Language and cognitive processes* 25(7). 1044–1098.

Bryzgunova, Elena A. 1980. Intonacija. In N. Ju. Shvedova (ed.), *Russkaja grammatika*, vol. 1, 96–122. Moskva: Nauka.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* New York: Harper and Row.

Cole, Jennifer, José I. Hualde, Caroline L. Smith, Christopher Eager, Timothy Mahrt & Ricardo N. de Souza. 2019. Sound, structure and meaning: The bases of prominence ratings in English, French and Spanish. *Journal of phonetics* 75. 113–147.

Frota, Sonia. 2002. The prosody of focus: A case-study with cross-linguistic implications. In Bernard Bel & Isabelle Marlien (eds.), *Speech Prosody 2002: Proceedings of the 1st international conference on speech prosody*, 315–318. Aix-en-Provence: Université de Provence. https://doi.org/10.21437/SpeechProsody.2002-64.

Ionin, Tania, Maria Goldshtein, Tatiana Luchkina & Sofya Styrina. 2023. Who did what to whom, and what did we already know? Word order and information structure in heritage and L2 Russian. *Linguistic approaches to bilingualism* 13(3). 343–371.

Jasinskaja, Katja. 2014. Information structure in Slavic. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 709–732. Oxford: Oxford University Press.

Kaiser, Elsi & John C. Trueswell. 2004. The role of discourse context in the processing of a flexible word-order language. *Cognition* 94(2). 113–147.

Katz, Jonah & Elisabeth Selkirk. 2011. Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language* 87(4). 771–816.

Kondrashova, Natalia. 1997. Generativnaja grammatika i problema svobodnogo porjadka slov. In Andrey A. Kibrik, Irina M. Kobozeva & Irina A. Sekerina (eds.), *Fundamental'nye napravlenija sovremennoj lingvistiki*, Moscow: Izdatel'stvo Moskovskogo Universiteta.

Lee, Goun, Dong-Jin Shin & Maria Teresa Martinez Garcia. 2019. Perception of lexical stress and sentence focus by Korean-speaking and Spanish-speaking L2 learners of English. *Language sciences* 72. 36–49.

Lobanova, Anna. 2011. The role of prominence scales for the disambiguation of grammatical functions in Russian. *Russian linguistics* 35(1). 125–142.

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

Luchkina, Tatiana & Jennifer S. Cole. 2021. Perception of word-level prominence in free word order language discourse. *Language and speech* 64(2). 381–412.

Luchkina, Tatiana, Tania Ionin & Maria Goldshtein. 2024. Acquisition of non-contrastive focus in Russian by adult English-dominant bilinguals. *Frontiers in Psychology* Volume 15 - 2024. https://doi.org/10.3389/fpsyg.2024.1363980.

Luchkina, Tatiana, Tania Ionin, Natalia Lysenko, Anastasia Stoops & Nadezhda Suvorkina. 2021. Evaluating the Russian language proficiency of bilingual and second language learners of Russian. *Languages* 6(2). 83.

Neeleman, Ad & Elena Titov. 2009. Focus, contrast, and stress in Russian. *Linguistic inquiry* 40(3). 514–524.

Ortega-Llebaria, Marta & Laura Colantoni. 2014. L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in second language acquisition* 32(2). 331–353.

Pierrehumbert, Janet & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 271–311. Cambridge, MA: MIT Press.

Pyykkönen, Pirita, Heiner Drenhaus & Matthew W. Crocker. 2011. Expectations of discourse salience: An ERP study of argument order preferences for ditransitive verbs. In *17th Annual conference on Architectures and Mechanisms for Language Processing (AMLaP)*, Paris, France.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural language semantics* 1(1). 75–116.

Sekerina, Irina. 2003. Scrambling and processing: Complexity, dependencies, and constraints. In Simin Karimi (ed.), *Word order and scrambling*, 301–324. Malden, MA: Blackwell.

Selkirk, Elizabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. In John Goldsmith (ed.), *The handbook of phonological theory*, 550–569. Cambridge, MA and Oxford: Blackwell.

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

Sorace, Antonella & Francesca Filiaci. 2006. Anaphora resolution in near-native speakers of Italian. *Second language research* 22(3). 339–368.

Swerts, Marc. 2007. Contrast and accent in Dutch and Romanian. *Journal of phonetics* 35(3). 380–397.

Swerts, Marc, Emiel Krahmer & Cinzia Avesani. 2002. Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of phonetics* 30(4). 629–654.

Wagner, Michael, Mara Breen, Edward Flemming, Stefanie Shattuck-Hufnagel & Edward A. Gibson. 2010. Prosodic effects of discourse salience and association with focus. Ms., MIT. https://dspace.mit.edu/handle/1721.1/71815.

White, Lydia. 2011. Second language acquisition at the interfaces. *Lingua* 121(4). 577–590.