2020

To Flip or Not to Flip?: Individual Cognitive Differences and L2 Russian Acquisition of Verbal Morphology in a Flipped Classroom

Katie M. Esser
University of Vermont

To Flip or Not to Flip?: Individual Cognitive Differences and L2 Russian Acquisition of Verbal Morphology in a Flipped Classroom

Katie Esser
Department of Romance Languages and Linguistics
University of Vermont
Honors Thesis
Spring 2020

Table of Contents

Acknowledgements **3**

Abstract **4**

Introduction **5**

Instructed Second Language Acquisition (ISLA) **6**

- ***Explicit Learning*** **6**
- ***The Flipped Language Classroom*** **8**

Individual Differences **10**

- ***Working Memory (WM)*** **10**
  - WM and SLA: The Connection Between WM Capacity and Language Proficiency 12
  - WM in the Acquisition of L2 Morphology 17
  - WM and Morphological Generalization 20

Morphological Processing **23**

- ***The Usage-Based Approach (UBA)*** **23**
- ***Factors in Morphological Processing*** **24**
- ***Morphological Processing of L2 Russian*** **26**

Russian Verbal Morphology **29**

- ***Foundational Studies*** **29**
- ***Russian Inflectional Paradigms*** **30**

Experimental Design **37**

- ***The Intact RUSS 002 Classroom*** **37**
- ***Participants*** **38**
- ***Procedure*** **38**

Hypotheses **46**

Results **50**

- ***Russian Verbal Morphology Test*** **50**
  - Comprehension Task 50
  - Written Production Task 50
- ***Cognitive Testing*** **52**

Discussion **55**

Conclusion **67**

Suggestions for Future Research **69**

References **71**

Appendices **77**

Appendix A: Gor & Chernigovskaya Studies ***77***

Appendix B: RUSS 001/002 and Mezhdu Nami ***80***

Mezhdu Nami Homework 80

Nemnogo o jazyke: Default -*aj* class conjugation table 81

Nemnogo o jazyke: Information on stress shifts in different conjugations 81

Appendix C: Raw Proficiency Test Scores ***82***

Appendix D: Written Production Generalizations ***87***

Appendix E: Russian Verbal Morphology Test ***90***

Acknowledgements

First, a thank you to my Russian professor, Kat “Katya” Scollins, who was very much the inspiration behind this thesis. I started taking Russian in my first year at UVM almost entirely on a whim. Katya’s love for Russian language and literature is absolutely infectious. A semester in, with lots of laughs surrounded by pages upon pages of grammar worksheets, I declared the major. Two and half years later, I spent a semester in St. Petersburg.

Katya’s teaching style is exciting and inspiring. Her classes and advising guided me toward the study of second language acquisition and teaching methodology. In prepping to become her Teaching Assistant for the year, we settled upon the idea that we could bridge the Teaching Assistantship with my thesis. It has been a great honor and pleasure to work alongside her in all of my four years, but particularly this year as I finish my undergraduate career.

To Guillermo Rodriguez, my thesis advisor and Linguistics professor: thank you! Your guidance in refining my topic, creating the proposal, submitting to IRB, editing countless drafts, et cetera, is more than appreciated. Bringing together my two majors through this project has been exhilarating, and I am more than thankful for your advising over the past year.

Credit is also due to Professor Diana Popa, who reviewed my self-designed test for thesis and agreed to join the committee. I thank you very much for taking time to review my work.

Finally, thanks to my parents, who are always there to give me reassurance and listen to me ramble about linguistics. They have always emphasized the importance of being an avid reader and continuing one’s education, which I fully believe is the drive behind my academic success. Thank you.

Abstract

The Elementary Russian curriculum at UVM is a flipped classroom, a relatively new approach to communicative language teaching, in which explicit grammar and vocabulary work is conducted at home and class time is reserved for communication between peers and the instructor. In this thesis, we measured the interactions between teaching methodology, learner cognitive capacity, and language proficiency in the acquisition of Russian as a second language (L2). As a means to investigate proficiency, we tested students' knowledge of the complex Russian conjugation pattern for present tense. Participants completed cognitive tests measuring working memory (WM) capacity, attention, multi-tasking capacity, and fluid intelligence. These variables were correlated with the proficiency results, which revealed significant relationships between WM and attention capacity. In the present study, WM and attention predict a learner's performance in the production of Russian verbal morphology.

Introduction

The study of second language acquisition (SLA) has exploded in recent decades. Applied linguists have devoted much time and effort into theorizing and supporting countless hypotheses about the cognitive abilities required in the successful acquisition of a second language in the language classroom. These studies are more often than not conducted in university-level language classrooms, in which students first take a proficiency test measuring knowledge of a specific grammatical feature in the target language. Subsequently, they complete one or more cognitive tests. Researchers are then able to run correlations and/or regression analyses between the two sets of tests to draw conclusions as to which cognitive variable(s) led to greater success in acquiring the second language (L2).

In this thesis, we take a very similar approach by administering 17 beginner students of L2 Russian a proficiency test assessing their knowledge of present-tense conjugation. This specific feature was chosen due to its high complexity (e.g. stress shifts, consonant mutations, etc. in some inflected forms). Students then took a series of five cognitive tests assessing their working memory (WM) capacity, attention, fluid intelligence, and multi-tasking capacity.

After running multiple correlations with our proficiency and cognitive variables, we uncovered a strong correlation between Russian proficiency, as measured by knowledge of Russian conjugation, and WM and attention.

The primary point of departure from previous work, however, was the consideration of teaching methodology as an influence on student performance. The classroom at the subject of this research is a flipped communicative classroom, meaning that traditional lecture material such as explicit grammar instruction is first assigned for homework and class time the next day is devoted to communicative practice to reinforce what was learned at home. Results from the multiple correlations were considered to put forth conclusions on the role of teaching methodology in student success in the L2 classroom, as well as some suggestions for future research and instructor intervention.

Instructed Second Language Acquisition (ISLA)

Explicit Learning

To understand the investigation carried out in this thesis, an understanding of the context of this particular L2 learning environment is necessary. Most second language classrooms are characterized by explicit learning. Ellis (2009) identified several defining characteristics of explicit knowledge:

(1) Conscious

(2) Declarative and fact-based

(3) Can be inaccurate and imprecise

(4) Is verbalizable

Characteristics (1) and (4) are similar in that both allow second language (L2) learners to explain issues of grammaticality in their own words. Not only are learners in explicit learning environments able to tell whether an utterance is grammatical or ungrammatical, but they also have awareness of the rules that they are applying to make this judgement and can verbalize them.

While learners have conscious awareness of their L2 explicit knowledge, it is important to consider (3), as the rules that learners acquire or extract from instruction may be incomplete and inaccurate. Ellis (2009) uses an L2 English learner's grammaticality judgment of "*The policeman explained Wong the law" as an example. This learner justified the ungrammaticality by stating that proper nouns cannot follow "explain." While the learner has arrived at a hypothesis, albeit erroneous, that the deems the sentence as grammatical, they have, of course, not developed a *full* understanding of the restrictions on that construction (i.e. it is ungrammatical because it lacks the dative "to PROPER NOUN" construction). As L2 acquisition proceeds, however, learners are able to refine these rules to a higher accuracy and completion.

Finally, (2) states that explicit knowledge is declarative and fact-based. L2 learners in instructed environments learn or extract rules that they can apply to their own language use and comprehension. These rules are unlike the systematic, implicit rules we know in our native language, and are verbalizable (Ellis, 2009). This type of knowledge is more similar to facts that we store, such as the year the Constitution was signed. On the other hand, an example of declarative knowledge would be an L2 Russian learner who knows that feminine nouns in the nominative case typically end in *-a* or *-ya.*

Much effort has been put into investigating the role of individual differences in explicit language-learning conditions, with varied results. For example, Tagarelli, Borges-Mota, & Rebuschat (2011) examined L1 English speakers' ability to judge the grammaticality of sentences in a semi-artificial language featuring English lexical items and German syntax. The performance of participants in the explicit learning condition was predicted by WM *only* on grammatical items in the grammaticality judgement task. The authors assert that although WM effects did not extend to the whole test, individuals with higher WM capacities may have been able to better absorb positive evidence, or grammatical information from the input they were afforded. Additionally, through post-test interviews, they found that the explicit knowledge derived from the exposure session that learners were able to verbalize was incomplete[1] (Tagarelli et al, 2011), corroborating Ellis' (2009) claims regarding the nature of explicit knowledge. On the other hand, Pawlak & Biedrón (2019), in a study assessing the production and comprehension abilities of upper-intermediate to advanced L2 English speakers, found a strong correlation between verbal WM and productive explicit knowledge. Learners with higher verbal WM scores were significantly more successful in orthographically producing correct English passive verbal forms than those with lower scores. Verbal WM also had strong, significant effects on receptive explicit knowledge (comprehension ability), although not as significant (Pawlak & Biedrón, 2019). As current evidence is inconclusive on the true relationship between explicit learning and individual differences (e.g. WM), further investigation is necessary. This is particularly pertinent as teaching methodology continues to evolve. Language teaching professionals are increasingly adopting more communicative methods that blend explicit learning with opportunity for meaningful production practice. Many variations exist in the currently established methods, including the more novel flipped classroom approach.

The Flipped Language Classroom

The flipped classroom (FC) approach was pioneered in the 1990s by a Harvard physics instructor who inverted his class structure by assigning lectures and readings for homework in order to devote class time for traditional "homework" activities. Since then, this approach has

[1] E.g. "verbs can appear at the end of sentences", without mentioning that this is only grammatical for subordinate clauses.

spread among the disciplines. In these classrooms, instructors are transformed from knowledge "transmitters" to facilitators, and students become active participants, increasing interaction and inquiry. In L2 FC contexts, students and instructors have more opportunities to communicate, rather than spending time on traditional grammar drills (Correa, 2015). Explicit grammar learning is assigned for homework, allowing students to process information at their own pace, which in turn increases their depth of processing and retention. In class the following day, this knowledge is bootstrapped through communicative activities that increase participation and L2 knowledge development, thus synthesizing explicit learning and communicative teaching methodology (Prefume, 2015). In addition to making the classroom more interactive, students in FCs demonstrate significantly greater levels of confidence and competence, which are crucial affective variables in learning an L2 (Tonkin, Page, & Forsey, 2019).

While this pedagogical approach is still novel to the field, several findings in the study of individual differences (see Individual Differences section) suggest that the pre-practice grammar lesson condition characteristic of the FC approach may equalize differences among students in a language classroom. For example, WM capacity effects seemed to level off for participants in an experimental L2 Latin study when they were exposed to explicit grammar instruction before completing a language task (Sanz, Lin, Lado, Stafford, & Bowden, 2014). In Li, Ellis, & Zhu's (2019) study, eighth grade L2 English learners were placed in a variety of different learning conditions to assess WM effects. In the classrooms in which grammar instruction was delivered before performing communicative tasks and no feedback was provided, WM had no effects. However, it is important to note that WM *did* have effects when both pre-task instruction and in-task feedback were delivered, as learners were supposed to balance language input and corrective feedback simultaneously (Li et al., 2019). In many FCs, students self-direct their pre-lesson explicit knowledge and are then given feedback in class during communicative practice. Due to these seemingly contradictory findings, we attempt to assess the effects of WM, in addition to several other cognitive variables, in conjunction with the FC in this thesis.

Individual Differences

Working Memory (WM)

The contemporary study of working memory (WM) was spearheaded by Alan Baddeley and Graham Hitch (1974), who sought a way to better describe the functionality and limits of this short-term memory capacity. In their pioneering research, Baddeley and Hitch designed experiments in which participants were given a processing task (listening comprehension, verbal reasoning, etc.) while also performing a memory task in which they committed to memory a string of numbers for immediate recall. They found that individuals vary in their capacity to complete these dual tasks and that this variation correlated with other cognitive capacities.

Baddeley continued to develop the original model and today his multi-component WM model is the framework adopted by a vast majority of researchers interested in the connection between memory and language. The current version presents memory capacity as a construct divided into four components:

(5) Central executive: the site of attention allocation, control and monitoring

(6) Phonological loop: a small buffer for the storage and rehearsal of verbal information

(7) Visuospatial sketchpad: a buffer for the processing of images, shapes, and locations

(8) Episodic buffer: this component brings together information from the slave systems (phonological loop and visuospatial sketchpad) and deals with integration into long-term memory (LTM; Li, 2017)

Arguably the most important component, the central executive, operates and controls other WM processes (6-8). Baddeley and Hitch (1974) considered the central executive as an intermediary between the phonological loop and the visuospatial sketchpad (6, 7). Current WM research also suggests that the central executive updates information and allows for multi-tasking, thanks to its attentional control function (Wen, 2016).

As mentioned above, the phonological loop is the site of phonological storage and articulatory rehearsal. This component handles acoustic input and it has been mostly implicated in the acquisition of vocabulary in an L2 (Gathercole & Papagno, 1998). The visuospatial sketchpad is also a site for storage and rehearsal, albeit for information in the visual and spatial modalities (Baddeley, 2007) and has not had a relevant role in language acquisition related research.

The episodic buffer was introduced as an interface between the three other components of WM (5-7) and LTM (Baddeley, 2000). The episodic buffer is accessed through conscious awareness. Unlike LTM, the episodic buffer and WM as a whole are temporary, but when the episodic buffer allows for information to be transferred to and encoded in LTM, long-term learning and retrieval become possible, which is crucial in the acquisition of a second language (Baddeley, 2007).

There are two characteristics of WM that determine an individual's capacity. On one hand, the amount, or availability, of information that can be stored for immediate recall is limited but varies between individuals. On the other hand, accessibility is the speed at which information can be recalled. An individual with high availability and accessibility, therefore, has a high WM capacity.

WM has traditionally been assessed through two instruments: simple or complex span tasks. Simple span tasks require individuals to store and rehearse information without additional interference, measuring storage capacity *only*, while complex span tasks require them to store and rehearse information while completing an additional task, measuring *both* storage and processing capacity. These complex tests require processing through a distracting task such as solving equations (Automated Operation Span Task), reading sentences for meaning/grammaticality (Reading Span Task), and comparing visual stimuli (Symmetry Span Task), and, at the same time, committing to memory sequences of letters, words or visual arrays (Linck, Osthus, Koeth & Bunting, 2014). In this thesis, the Operation Span Task and the Symmetry Span Task are used to assess WM capacity among our beginner L2 Russian learners, in an effort to assess both verbal/arithmetic and visual abilities.

WM and SLA: The Connection Between WM Capacity and Language Proficiency

In recent years, WM research has focused heavily on the connections between the theoretical model and how it operates in second language learning. Edward Wen (2016) argues that WM in second language acquisition (SLA) should be sub-divided into phonological and executive WM, with each serving a different function. Phonological WM, connected to the phonological loop of Baddeley's model, is the component controlling sound processing and is said to aid in acquisition of vocabulary, formulaic language, and morphosyntax. Executive WM, similar to the central executive in Baddeley's model, controls attention and monitors L2 performance and processing in all modalities (listening, speaking, reading, and writing). It has also been suggested that it is particularly at the beginning stages of SLA that WM has a greater effect overall (Serafini & Sanz, 2016). Through additional investigations, researchers have determined that phonological WM plays a greater role in increasing proficiency at the beginning stage, while executive WM is more important at the intermediate and advanced levels of L2 proficiency (Wen, 2016).

Researchers have also recently explored the role of WM in foreign language aptitude (FLA), or the specific "talent" for success in learning an L2 some individuals present (see Wen & Skehan, 2011; Miyake & Friedman, 1998). WM has been found to be a significant factor in accounting for variance in beginner L2 proficiency, as measured by test scores and GPA (Linck & Weiss, 2011; Miyake and Friedman, 1998). Therefore, it is claimed that WM can predict L2 learning, meaning that learners with greater WM capacities typically exhibit greater improvements in their L2 proficiency. Because processing an L2 is thought to be a demanding cognitive activity for adults (like solving complex math equations or theorems), it is reasonable to believe that a greater WM capacity, with greater attentional control, would predicate higher ultimate attainment in an L2 (Linck & Weiss, 2011).

Although there are promising results regarding the connection between WM and SLA, until recently, there has not been consensus on the extent of the association between the two. In fact, some studies over the past few decades have demonstrated inconsistencies, due to variations in WM capacity and proficiency testing measures, as well as other study design features such as population size.

A recent meta-analysis (Linck et al, 2014) included 79 studies to determine the relationship between WM and L2 performance (measured by processing and proficiency). Through this analysis, a positive correlation was found between WM and L2 attainment. Additionally, researchers concluded that the central executive (measured through complex span tasks), which controls allocation of resources for storage and processing, is strongly associated with L2 progress. This further suggests that executive control is the strongest predictor of L2 success, as processing, attentional control, and multi-tasking are all simultaneously involved in use of the L2.

Research from Weissheimer & Mota (2009) and Martin & Ellis (2012), both included in Linck et al's (2014) meta-analysis, further demonstrate a positive correlational relationship between WM capacity and L2 attainment.

In Weissheimer & Mota (2009), adult English as a Foreign Language (EFL) learners were administered a Speaking Span test (to assess WM capacity) and a Speech Generation task (to assess L2 proficiency) twice over an 8-week period. In the WM test, participants were shown a sequence of words, spaced apart by a second. At the end of each set, participants were asked to create sentences in English containing words from the sequences just shown. A participant's speaking span, and broadly, their WM capacity, was calculated based on the number of grammatical English sentences generated, weighted based on the number of words used from each set. In the proficiency test, participants were shown a single picture and asked to describe a story based on the visual stimuli. L2 proficiency was determined through the participant's fluency, accuracy, and complexity of speech in the speech generation task. Ultimately, it was found that WM capacity and speaker accuracy and complexity in L2 English speech production were positively correlated. L2 learners with high WM capacities were able to produce speech that was more grammatically accurate, more lexically complex, and more structurally complex than learners with lower WM capacities. However, there was no significant effect found between WM capacity and fluency, perhaps due to the focused attention on accuracy and complexity. When it came to the development of L2 proficiency over time, fluency and complexity measures from the Speaking Span test were significant predictors at the conclusion of the 8-week period of testing, likely due to the proceduralization or automatization of explicit knowledge learned from classes in the L2, leading to more fluent and complex speech in individuals with high WM capacities (Weissheimer & Mota, 2009).

In a more recent study by Martin & Ellis (2012), the effect of phonological short-term memory (PSTM) and WM were analyzed regarding their relationship to vocabulary and grammar learning, respectively. To measure PSTM, a non-word repetition and a non-word recognition test were administered to participants. To measure WM capacity, participants took a Listening Span test, in which they listened to sentences in English and judged each sentence as grammatical or ungrammatical, and then were tasked to recall the final word of each sentence. The Listening Span score was calculated as the number of words correctly recalled in the correct order. After the two tests assessing individual differences, the participants were tasked to learn singular forms of words and sentences in an artificial language. After this exposure session, participants were tested on their production and comprehension of sentences that contained plural forms of the learned, non-plural vocabulary items. Both PSTM and WM measures were found to contribute to the learning of vocabulary. Specifically, PSTM helped explain 14% of variance between individual final vocabulary mastery, while WM as a whole explained 10% of the variance. On the other hand, WM was found to be positively correlated with grammar proficiency in production, while PSTM had a much weaker effect. Ultimately, WM capacity is a stronger predictor of the development of grammatical proficiency in an L2 because unlike PSTM, WM entails storage *and* processing, allowing learners to infer patterns from input to apply to future instances in production and comprehension (Martin & Ellis, 2012).

The meta-analysis (Linck et al, 2014) as well as the two studies described above, provide further evidence of the positive relationship between WM capacity and L2 proficiency. Specifically, it has been determined that speech production (as defined by fluency and complexity) is significantly positively correlated with WM capacity in L2 learners, as evidenced by Weissheimer & Mota (2009). Individuals with high WM capacities, therefore, are more successful in producing fluid speech with more complex vocabulary and grammar structures. These more advanced production skills are characteristic of the development of morphosyntax in the L2, as WM aids learners in extracting patterns and memorizing information, which provides for easier, more fluid future access. Martin & Ellis (2012) have shown that high WM capacity predicts grammatical proficiency in an L2. These findings are crucial, as they suggest that WM may be one of the key factors in explaining variation in L2 performance among students, whether it be through specific measures of speech production or through mastery of grammatical forms. Overall, WM stands as a significant predictor of success at L2 morphosyntactic acquisition.

In the present study, complex WM span tasks will be used to measure WM capacity. As discussed above, these complex tasks will assess processing, attentional control, and multi-tasking, all of which are involved in the different modalities of L2 use. To correlate these measures with language knowledge, a proficiency test will be administered, measuring accuracy in students' listening comprehension and written production of Russian verbal morphology. To successfully master the patterns of Russian conjugation, students need to process language, control attention, and multi-task as they assess input for function and meaning simultaneously. Since we utilize morphology as a means of assessing L2 proficiency, let us now focus on findings specific to the connections between this component of language and WM.

WM in the Acquisition of L2 Morphology

The study of the acquisition of L2 morphological markers has revealed that WM is a significant predictor of success in the L2 classroom. For example, learners with high WM capacity, and more specifically, inhibitory control (a component of executive function), were better at acquiring mastery of the morphosyntax of a simple artificial grammar with only 12 nouns and 4 verbs in both comprehension and production modalities (Kapa & Colombo, 2014). The adult L2 learners in this study with high WM capacity were able to mitigate or even theoretically by-pass the common strategy of translation from the L1 to the L2 in their second language. The ability to inhibit the initial access of L1 vocabulary is important in advancing L2 proficiency (Kapa & Colomobo, 2014). High-capacity individuals have also been shown to be more sensitive to morphological agreement violations of gender and number in English and other languages (see Sagarra & Herschensohn's (2010) L2 Spanish study). In these morphological agreement studies, a difference in cognitive load has been evidenced between gender and number agreement, with number agreement being less cognitively taxing (Sagarra & Herschensohn, 2010). The understanding that some aspects of morphology are harder to master than others is crucial to comprehending morphological processing (see the Morphological Processing section). The Spanish study, in particular, demonstrated that beginning learners of an L2, especially those with lower WM capacities, are not sensitive to agreement violations, even when presented with both positive and negative evidence. To improve grammatical sensitivity and therefore advance proficiency, Sagarra & Herschensohn (2010) suggest that learners need a greater amount of exposure to the L2. Researchers assert that a high WM capacity, with greater ability to recognize patterns and to avert attention away from distractors, allows learners to master L2 morphology more effectively.

While these and other studies indicate promising results for the connection between WM and L2 success, there is also evidence that suggests that this connection may be weak or even null. For example, McDonough and Trofimovich's 2016 study explored L2 Esperanto learning of a transitive morphological structure. In this construction, a single morpheme -*n* is added to the end of nouns to indicate that the word is an object. Because the morphology indicates function, word order can be flexible in Esperanto, meaning the object of the sentence can be placed before *or* after the verb, unlike in the participants' L1 Thai. For example, in (9), the Esperanto sentence follows the standard SVO (subject-verb-object) order, while in (10), it follows the atypical OVS order (object-verb-subject).

(9) *kato pelas cevalon.*

cat chases horseACC

Cat chases horse.

(10) *cevalon pelas kato.*

horseACC chases cat

Cat chases horse.

In the first part of the experiment, the exposure session, participants were shown two pictures while they listened to Esperanto sentences in the standard SVO or atypical OVS word order and were then asked to choose the correct picture. In this exposure session, participants could primarily rely on their lexical knowledge in connection with the picture stimuli, however, during the test, they solely relied on their functional knowledge of the *-n* suffix. Throughout the experiment, participants were never explicitly told to search for clues affording grammatical meaning. Ultimately, McDonough & Trofimovich found that WM was *not* a significant predictor of pattern recognition or L2 success. They claim that the effects of WM on L2 proficiency are typically observed only when individuals are explicitly told to pay attention to elements of the input during the exposure session. Because the exposure and testing conditions of the experiment were both designed to be implicit, this may explain the lack of a relationship between WM and L2 proficiency among the participants. Furthermore, exit questionnaires indicated that very few participants could recall explicit information about the grammar they were taught and tested on. The null effect of WM on L2 learning in this study suggests that the participants, therefore, did not seem to derive any explicit rules from the learning session (McDonough & Trofimovich, 2016).

In contrast to the findings described above, the subject population of this thesis has experienced explicit instruction on the topic of verbal morphology over the course of a full, 4-credit semester course, albeit at home (see Flipped Classroom section). Due to this extended explicit exposure, it is expected that these students will have gained explicit knowledge regarding morphological rules for verbal inflection in Russian, and therefore, WM capacity is expected to be a significant predictor of their command of L2 Russian morphology.

WM and Morphological Generalization

Another interesting finding is that WM has also been shown to predict the generalization of patterns to novel items, but not the application of inflection to known (previously encountered) items. For example, Kempe, Brooks, & Kharkhurin (2010) found that L2 learners of the Russian gender system were fairly successful in generalizing gender to novel items, with individual cognitive differences significantly predicting their variability in performance. Kempe et al. (2010) exposed absolute beginners of L2 Russian to masculine and feminine diminutive nouns in simple pairs, for example:

(11) *kukla* [doll-FEM]

(12) *kukolka* [doll-FEM_DIM]

Students who learned Russian gender morphology through diminutives before simple nouns typically demonstrated a greater mastery of the gender system, as all diminutives have *transparent* gender morphology, while simple nouns can be transparent (11-12, 15-16) *or* opaque (13-14) regarding gender morphological encoding.

This specific feature of Russian grammar is similar to the verbal morphology targeted in this thesis, as some nouns in Russian can carry gender morphology that could suggest more than one grammatical category. For example, *opaque* nouns in Russian often end in palatalized consonants yet belong to different grammatical genders (13-14):

(13) *pech'* [oven-FEM]

(14) *pen'* [stump-MAS]

(15) *pech'ka* [oven-FEM_DIM]

(16) *pen'chik* [stump-MAS_DIM]

The only way to infer the gender of nouns like these is through the natural tallying of statistical probabilities. Learners collect language sample data through exposure and are able to extract probabilities such as: it is more common for a noun ending in a palatalized consonant to be *masculine* or *feminine* (as opposed to neuter); or through the diminutive form of the same noun, which always has *transparent* gender morphology, with feminine nouns ending in -*a* and its allomorphs and masculine nouns ending in consonants (15-16).

During the exposure sessions, learners were shown picture stimuli while listening to a Russian noun. They were then asked to choose whether the noun belonged in the first or second gender category (avoiding "masculine" and "feminine" to prevent interference from knowledge of other gendered languages with similar gender morphology). Following several learning sessions, the participants completed a test similar to the learning task including diminutive and bare nouns with *transparent* gender morphology, some of which were already known and some of which were novel, as well as simple nouns with *opaque* gender morphology, known and novel.

Kempe and colleagues found that non-verbal intelligence (measured through the Cattell CFIT, similar to Raven's Progressive Matrices) was a more accurate predictor of success in relation to performance on learned and novel diminutives, where the morphological ending and gender of all diminutives is transparent. On these items, learners had to rely on their extraction and generalization of morphological patterns (created from exposure in the learning sessions). In general, learners with higher non-verbal intelligence were more successful in predicting gender category for all diminutives. On the other hand, the researchers discovered that individuals with higher WM capacities (measured through a Reading Span Task) were more accurate in the gender categorization of novel opaque nouns. These more successful participants were able to create and access a link between gender and word endings (generally -*a* for feminine, -consonant for masculine). The successful learners seemed to memorize connections between gender and endings and showed high accuracy in gender category recognition. These findings are interesting to consider as the experimental design was implicit both in that learners in the study did not know that novel items would be tested and they also did not make explicit judgements during testing (when selecting noun categories, learners chose from a red or green button instead of the more explicit "feminine" and "masculine" category). While this condition did allow to researchers to control for knowledge of other languages with similar gender morphology (-*a* is a common ending for female nouns in Spanish as well, for example), it also failed to replicate the typical explicit nature of language instruction. Although Kempe and colleagues' results are counter to those from McDonough and Trofimovich (2016), it is important to note that the explicit training and testing condition is crucial in replicating the average SLA environment, and thus more readily generalizable to L2-learning populations.

While previous studies on Russian *verbal* morphology have not yet assessed the effect of WM capacity (Gor & Chernigovskaya 2000, 2001, 2003, 2005; Tkachenko & Chernigovskaya, 2010), the beginner Russian students, who took part in previous studies, were successful at generalizing conjugation patterns to novel verbal stimuli, perhaps due to their prolonged explicit instruction experienced in their college-level Russian classroom, allowing them to extract and later generalize patterns that were afforded to them through input over the previous semester. As such, the proficiency test designed for our own research includes verbal items that are novel to the participants in order to assess whether memory capacity is correlated with accurate morphological deployment for learned and novel items.

Morphological Processing

The Usage-Based Approach (UBA)

Many researchers in SLA have more recently focused on a relatively novel theory in L2 morphological processing that moves away from the acquisition of abstract representations, the usage-based approach (UBA) (see Ellis, 2002, 2008; Tyler, 2010, Tkachenko & Chernigovskaya, 2010). Unlike nativist theories of language acquisition (Chomsky, 1981), the UBA asserts that humans are *not* innately endowed with domain-specific linguistic representations (or rules). Instead, we construct language through *domain-general* cognitive skills (Behrens, 2009), such as statistical learning (transitioned probabilities at all levels of linguistic analysis, analogical reasoning, grounded cognition, etc.). These general cognitive skills allow humans to unconsciously tally frequencies at all levels of representation in language (sounds, morphemes, verbs, nouns, grammatical constructions, etc.) from the abundance of positive evidence, or language input (Langacker, 1987). In other words, the UBA claims that the development and processing of language relies heavily on input frequencies (Ellis, 2002).

Those who support the UBA claim that, while there are differences between L1 acquisition and SLA, both modalities of language learning involve the abstraction of grammatical patterns through extended exposure to language input, and that these abstractions are sensitive to frequencies. The abstractions encode form-function mappings, which can later be applied to future comprehension or production tasks (Ellis, 2002). For example, English speakers have a mapping between *-s* and the 3rd person plural form of regular verbs, which they can spontaneously apply to all verbs that follow this pattern. The idea that when individuals learn a language, the patterns extracted from the input can later be generalized to unique and more complex utterances is key in understanding L2 acquisition under the UBA. In other words, new utterances *emerge* from these patterns derived from human interaction (Behrens, 2009; Ellis, 1998).

Factors in Morphological Processing

Taking this exposure-based view into account, in the realm of morphosyntactic processing, several variables come into play that can hinder fluid comprehension and/or production. One, for example, is the issue of morphological transparency, also known as cues. Opaque patterns in any language are more difficult to master because their forms or endings are seen as “irregular” or “exceptions” that need to be memorized rather than generated from abstracted patterns.

Another issue relates to the handling of irregularity, which has been explored through the English past-tense paradigm. L1 English speakers can produce high-frequency irregular verbs much more rapidly than low-frequency irregular verbs. When tasked to produce regular verbs of varying frequencies, L1 speech production does not demonstrate frequency effects, likely due to the power law of learning, as the accurate production of regular verbs is close to asymptote (Ellis & Schmidt, 1997). Processing of irregulars and regulars in an L2, however, is slightly different. In the beginning stages of acquisition, learners demonstrate frequency effects for regular *and* irregular forms. As proficiency increases, frequency effects on regular forms decrease whereas they remain the same for irregular forms (Ellis & Schmidt, 1997). In other words, the effect of frequencies on the acquisition, production, and comprehension of regular word forms decreases when learners progress to higher levels of language proficiency, while the effect remains stagnant for irregular word forms. Irregular forms, therefore, can pose difficulty at all stages of L2 proficiency, as they remain sensitive to frequencies afforded to learners from the input. Neurological evidence has been uncovered to this effect in relation to English past tense forms, as advancing proficiency demonstrates less neurological activation for irregular verb forms and insignificant activation for regular verb forms (Roncaglia-Denissen & Kotz, 2015).

Additionally, salience is crucial in the acquisition of L2 morphology. Grammatical morphemes (e.g. past tense -ed and present progressive -ing in English) are often bound by inflection, phonologically unstressed, and difficult to explicitly notice. These morphemes can also be redundant when lexical items contain the same functional meaning (Ellis, 2008). Examples (17) and (18) from the Russian language demonstrate this redundancy. In (17), the verbal morpheme is redundant because (18) provides the listener with a clue as to the person and number of the verb.

(17) *-jut* 3rd person plural suffix

(18) *oni* “they”

Therefore, it is easy for L2 learners to process the lexical item and gain no further information from the morphological suffix in the verb-bound inflection (Ellis, 2008).

Type and token frequencies are another important issue to note when examining the acquisition of L2 verbal morphology. Type frequencies refer to the size of a specific class of words (Tkachenko & Chernigovskaya, 2010). For example, in Russian the largest verbal class by type frequency is the *-aj* stem because that category contains more individual verbs than any other verbal class. Token frequency, however, refers to the frequency of a specific word (Tkachenko & Chernigovskaya, 2010). For example, a verb such as *dumat'* "to think" has high token frequency because it is encountered very frequently in Russian speech.

Morphological Processing of L2 Russian

The processing and acquisition of L2 Russian morphology and its relationship to individual differences has been explored in several ways, namely through gender and case paradigms. For example, Kempe & Brooks (2008) investigated L2 morphology acquisition through a miniaturized version of Russian, in which ab-initio learners were tasked to learn two grammatical genders and three cases for 24 nouns exclusively in the aural modality. In their first experiment, Kempe & Brooks tested half of the learners on nouns with transparent gender markings in the masculine and feminine while in the second test they tested the other half on opaque nouns in both the masculine and feminine (see examples 7-12). At the conclusion of language testing, Kempe & Brooks administered a WM and non-verbal intelligence test to correlate participants' L2 success with individual differences.

In the experiments, through several learning sessions, subjects were provided with picture stimuli of singular objects to prompt specific grammatical cases (Figure 1).

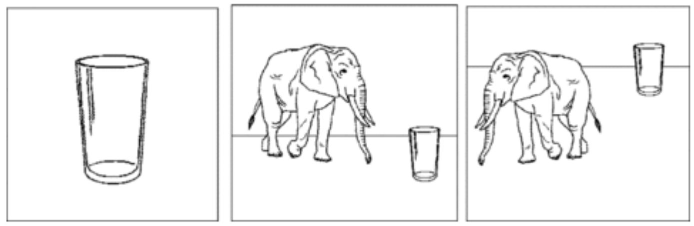


Figure 1: Test stimuli in Kempe & Brooks, 2008

All participants were assessed on their comprehension of vocabulary and case and production of case endings.

Ultimately, the subjects who learned exclusively transparent nouns were more successful in acquiring the simplified Russian grammar and lexicon. However, they produced gender suffixes more accurately than case suffixes when required. Individuals in this experiment with high nonverbal intelligence were more successful in producing correct endings for novel nouns, indicating that they were more successful at generalizing.

In the second experiment, individual success in learning opaque nouns was significantly associated with the storage capacity of their WM. The connection between success with opaque nouns and WM storage capacity appears justified under the model that Kempe et al. endorse, which posits that opaque forms are learned through rote memorization. Secondly, after prolonged exposure, the participants in the second experiment with higher non-verbal intelligence were able to extract suffix patterns and apply case inflections to novel items. Again, this falls in line the researchers' model, as intelligence is thought to influence the learning of transparent forms due to the construction of patterns from input (Kempe & Brooks, 2008).

However, in both experiments, most learners showed generally poor performance on novel stimuli, a corroborated finding from previous studies on L2 Russian acquisition (Kempe &

Brooks, 2005). This finding, could, however, be triggered by the testing of both gender and case paradigms. Testing both paradigms means that learners are expected to derive not just one, but *two* form-based meanings from each morphological form on a single noun.[2]

Analyzing these results, it is evident that there may be a benefit from teaching an L2 in a way that places more weight on transparent versus opaque morphological items (see also the diminutive study by Kempe, Brooks, & Kharkhurin, 2010). Although students taught in this way have less exposure to the more cognitively demanding stimuli, they may be better equipped to make accurate generalizations of transparent morphology, which follows a regular paradigm. Perhaps learners with a more balanced exposure to transparent and opaque morphology do not form as strong of a connection between form and meaning of both transparent and opaque morphology than if these learners were exposed to a more imbalanced input, in which transparent morphology is heard and used more frequently. Conveniently, it is also the case that lexical items containing transparent morphology are typically more frequent and productive in a language (see the Russian Verbal Morphology section). In the Experimental Design section, we will describe the balance between each verbal class represented in the input from the textbook homework and classroom activities and whether it maps to the balance represented in authentic, native input.

Although the Russian gender system is indeed complex, Russian verbal morphology is a much more nuanced system that is affected by regularity, frequency, and complexity. Due to

[2] Additionally, Russian has case markings that are highly opaque, for example, the -*u* ending which is found in the feminine accusative and masculine dative.

the accumulation of these variables, researchers have posited a model that goes beyond the transparent-opaque distinction characteristic of the Russian gender paradigm.

Russian Verbal Morphology

Foundational Studies

Beginning in 2000, Kira Gor and Tatiana Chernigovskaya set out to build upon the body of research on verbal morphological processing, previously conducted with Norwegian, Icelandic, and Italian speakers (Matcovich, 2001; Ragnasdóttir, Simonsen, and Plunkett 1997). The most widely accepted processing model for morphological information is the dual processing approach, which postulates that regular forms are processed by rule-based mechanisms, while irregular forms are processed through associative memory (Gor & Chernigovskaya, 2000). Evidence from these previous studies indicated that the dual-system approach to verbal processing does not necessarily fit morphologically rich languages, which have gradations of "regularity" and complexity. Using Russian verbal morphology as a vehicle, the researchers explored several aspects of processing:

(19) Default pattern and generalizations

(20) Type frequency and complexity factors

(21) Role of morphological cues

(22) Application of morphological "sets"

In addition to examining the former components of morphological processing, Gor and Chernigovskaya compared this process among their two experimental populations (children and adults) to account for:

(23) Similarities and differences between L1 and L2 processing

(24) Factors accounting for differences

(25) Role of type frequency and complexity in L1 and L2 processing

In a series of studies, Gor & Chernigovskaya have explored various topics related to Russian verbal morphology (see Appendix A).

Russian Inflectional Paradigms

Before understanding the findings of these studies, it is pertinent to explain the Russian conjugation system. The most widely accepted and utilized description of the Russian conjugation system was first described by linguist Roman Jakobson (1948). There are 11 verb classes, each with a unique suffix that predicates the conjugation type (1st or 2nd) and morphological alternations; including consonant mutation, stress shift, and suffix alternation (Gor & Chernigovskaya, 2004). While the progression of Gor, Chernigovskaya, and Tkachenko's research includes most verb classes, there are four classes that are the primary focus of most experiments: *-aj, -a, -i,* and (*-ova*). These four verb classes vary in productivity, type frequency, and morphological complexity. Productivity entails an affix's ability to apply widely or narrowly. Affixes with high productivity occur very frequently in the language. Type frequency refers to the size of a particular class, for example, *-aj* is the largest class by type frequency because a sizable majority of Russian verbs fall under this class (Gor & Chernigovskaya, 2003). Type frequency is often not representative of the target language in an L2 environment, as learning conditions are controlled. Thus, L2 learners' verb classes may be more equally represented versus L1 speakers', who have access to authentic native input which may be imbalanced toward a specific, more productive verbal class (Gor & Chernigovskaya, 2000). Morphological complexity for each verbal class was determined through the number and type of rules each class is subjected to (Gor & Chernigovskaya, 2000). Figure 2 shows the characteristics of each verbal class' morphological complexity, including stress shift, consonant mutation, and more.

Class	Alternation (inf/past ~ present)	Conjug. type	Morpho-logical complexity	Type frequency/ Productivity	Examples (Inf/3pl/1sg)
-AJ-	a ~ aj	1	—	Very high/ Productive	igrá-t igráj-ut igráj-u
-A-	a ~ ∅	1	Consonant mutations, stress shifts	Low/ Unproductive	pisá-t píš-ut piš-ú
-I-	i ~ ∅	2	Consonant mutations, stress shifts	High/ Productive	nosí-t nós'-at noš-ú
-OVA-	ova ~ uj	1	Suffix alternation	Medium/ Productive	risová-t risúj-ut risúj-u

Figure 2: Russian verbal classes (from Tkachenko & Chernigovskaya, 2010)

As we can see, for example, the *-i* class, is complex because it has two morphological rules, consonant mutation and stress shift, that apply from the original stem into the conjugated form. In contrast, the *-aj* class is more straightforward, as it contains none of those complex morphological features.

In all stems for all verbal classes, automatic truncation takes place in at least one tense. In verbs with stems ending in vowels and inflections ending with vowels, the first vowel is truncated. The same process follows for verbs with stems and inflections ending in consonants.

Gor & Chernigovskaya (2000) demonstrate this effect with *-aj-* class verb *chitat'* "to read" and -*a*- class verb *pisat'* "to write":

(26) Stem: *chit-aj* + Inflection: *-l* = Past tense: *chital* "He read"

(27) Stem: *chit-aj* + Inflection: *-u* = 1st person singular: *chitaju* "I read"

(28) Stem: *pis-a* + Inflection: *-l* = Past tense: *pisal* "He wrote"

(29) Stem: *pis-a* + Inflection: *-u* = 1st person singular: *pishu* "I write"

In (26), the *-t* ending the stem causes the *-j* to be truncated before the *-l* consonantal past tense inflection. In (29), the *-a* is truncated and replaced by the 1st person singular inflection, *-u*. We will see later that automatic truncation is an important rule that L1 and L2 speakers deal with differently in experiments.

Additionally, it is paramount to note that the *-aj* stem is unrecoverable in its infinitive and past tense forms. (26) and (28) demonstrate this, as the two past tense forms of an *-aj* and an -*a* class verb contain the same word-final vowel, *a*. Unlike in its present tense counterparts, *chitat'* "to read" does not exhibit the characteristics of an *-aj* class verb in its infinitive and past tense forms. Therefore, we should expect that L2 learners would generalize the *-a* class stem to verbs like *chitat*. On the contrary, the interlanguage of beginner L2 Russian speakers exhibits the opposite effects, likely due to frequency effects. It is much more common that these speakers produce forms such as *pisaju* (unfortunately meaning "*I urinate*") for *pisat'* "to write", and not the correct 1st person singular form as seen in (29). Although the *-aj* stem is unrecoverable, it is much more typologically frequent in both L1 and L2 input and also does not exhibit the same degree of morphological complexity that *-a* class verbs do in their present tense inflections.

For the purposes of this study, we will only be concerned with the first three classes of verbs because RUSS 001 students only learn verbs belonging to the first three verbal classes.

The populations of all the studies reviewed include a beginner L2 group of 15-20 American university students and either an L1 group children (Gor & Chernigovskaya, 2003, 2004, 2005; Tkachenko & Chernigovskaya, 2010) or adult university students in St. Petersburg, Russia (Gor & Chernigovskaya 2000, 2001). The studies with child L1 populations focused on determining whether L2 processing of Russian verbal morphology is similar to that of a specific L1 age group. Conversely, studies with adult L2 populations focus on adult native processing rather than processing in developmental stages. This thesis, however, is only concerned with L2 processing. Appendix A shows a fully-fledged comparison of the experimental conditions of each study.

Gor and Chernigovskaya (2000, 2001) designed an experiment in which adult L1 speakers were tasked to conjugate 48 minimal pair[3] nonce verbs and adult beginner L2 speakers generated 48 learned verbs. Both groups were prompted for responses in the 3rd person plural and 1st person singular via oral dialogue:

Experimenter: Yesterday they _____. And what are they doing today?

Subject: Today they ______.

Experimenter: And you?

Subject: Today I ______.

(Gor & Chernigovskaya, 2001)

[3] The nonce verbs were minimal pairs because they differed in sound from real verbs by only a single phoneme.

In Gor & Chernigovskaya (2003, 2004) and Tkachenko & Chernigovskaya (2010), L2 learners were presented with an equal number of real and nonce verbs. Given the low proficiency of these learners, it may be surprising that they were asked to generate morphological markers for non-existent verbs. Nevertheless, the motivation behind this experimental condition is to determine L2 learners' ability to generalize inflectional paradigms, influenced by input frequencies (Tkachenko & Chernigovskaya, 2010). As the researchers postulate, if frequency effects are significant in regular verb processing, learners master conjugation through associative patterning (Gor & Chernigovskaya, 2003). After being afforded with sufficient input, they should be able to extend those patterns to novel items, or nonce verbs in this case. The L2 population of this thesis were presented with an equal number of real, known and nonce (unknown) verbs in accordance with these research studies.

The remaining studies only present real verbs to L2 learners, focusing on generalizations and their influences (Gor & Chernigovskaya, 2000) and the role of explicit instruction and input token together with type frequencies (Gor & Chernigovskaya, 2005).

This thesis focuses heavily on the role of a particular type of explicit instruction, and thus, Gor & Chernigovskaya's (2005) study will be in the target of the partial replication involved in this research.

A unique characteristic of the 2005 study's experimental design is the consideration of input frequencies in the instructed L2 environment. The researchers tallied the type frequencies of each verbal class from the two textbooks and workbooks used in the L2 population's first-year Russian program. The type frequencies used in the L2 classroom are roughly proportional to that of native Russian input, albeit not equivalent due to the reduced amount of exposure characteristic of a college-level language learning environment.

A similar experiment as described in the 2000-2010 studies was conducted, albeit delivered in both written and oral modalities.

An analysis of the results led to three main findings:

(30) L2 learners' processing is influenced by statistically tallied type frequencies. High frequency conjugation paradigms are more commonly generalized to other verb classes.

(31) L2 learners relied less on default conjugation paradigms due to input frequencies. The input afforded to this population is different than what is afforded to L1 speakers.

(32) L2 learners can apply explicitly taught rules successfully and produce native-like generalizations of default patterns to other verb classes.

Through these conclusions, Gor & Chernigovskaya (2005) suggest that, due to the fact that the L2 learners were able to perform at a native-like capacity, explicit instruction must have been a key element in their success. Participants in this experiment learned the features of Russian verbal morphology through a focus-on-form approach, which embeds deliberate attention to grammatical form into meaningful context (Long, 1991). While the intact classroom that is the subject of this research *does* utilize this pedagogical approach, it does so in a different manner (the FC approach). This thesis will partially replicate the 2005 study, with the addition of cognitive testing in an effort to determine the role of the FC as a teaching methodology.

In Tkachenko & Chernigovskaya (2010), the L2 learners demonstrated high rates of stem recognition for the *-aj* and *-i* verbal classes, the two most frequent and productive verbal classes in Russian. However, it is important to note that these learners generalized the highly productive, high type frequency -*aj* stem to *over* 25% of nonce stimuli (more than necessary, as the four verbal classes tested were given equal weight in the test stimuli). While -*a* class recognition was low, L2 learners were in fact more likely to generalize the -*a* class stem to novel stimuli even though the -*a* class stem makes up a very small proportion of the afforded input. Tkachenko and Chernigovskaya suggest that, in this instance, the linguistic environment of the L2 learners has exerted more of an influence than the type frequency of the stem. In their instructed, focus-on-form classroom, the L2 learners were afforded explicit instruction on the "irregular" conjugation paradigm of the -*a* class stem, which in turn increased the salience of any instance of the stem. Therefore, the L2 learners were able to successfully generalize "irregular" patterns, likely due to the influence of the explicit instructed environment in which they were acquiring Russian.

Although the aforementioned studies brought forth important information on how L2 learners of Russian process the complex system of verbal morphology, no work has yet been done on how individual cognitive differences influence learners' acquisition of this particular element of morphosyntax. Additionally, although all of the studies on this feature were conducted in classroom that taught through explicit instruction, teaching methodology was not considered in the analysis of results.

In this thesis, a Verbal Morphology Test derived from Kempe & Brooks (2008) and Tkachenko & Chernigovskaya (2010) was deployed to assess beginner L2 Russian students' knowledge of the verbal conjugation paradigm. To build upon the established body of work, students then completed a series of cognitive tests, assessing WM capacity, attention, intelligence, and multi-tasking capacity. In performing this experiment, two research questions were taken into account:

(33) Is there a specific cognitive profile that is more or less likely to excel in the intact RUSS 002 classroom? For example, does higher WM capacity predict higher performance on Russian verbal morphology?

(34) What is the effect of the flipped classroom approach on student performance?

Experimental Design

The Intact RUSS 002 Classroom

The RUSS 001/002 curriculum at the University of Vermont (UVM) uses the communicative online textbook *Mezhdu Nami* (DeBenedette, Comer, Smyslova, & Perkins, 2019). Students are assigned several pages of homework for each class, out of a workbook designed by the textbook authors. Each worksheet corresponds to a unit and sub-unit number containing dialogues from the online textbook. Students listen to the dialogues and read along, noting new, bolded vocabulary (Appendix B). Then, they are expected to read explicit explanations of any new terms or grammatical concepts in *Nemnogo o jazyke* ("A little about the language", Appendix B). The dialogue in addition to *Nemnogo o jazyke* is meant to inform their completion of the homework assignment and prepare them for class the next day. The following day's in-class activities are communicative and reinforce topics learned at home the previous night.

The default, most productive verbal class, -*aj*, is explicitly taught in the online textbook and a pattern is provided for students to apply as a rule to other similar verbs (Appendix B).

However, for more complex inflection paradigms, such as *-a*, no rule or pattern is provided. Students are taught that stress shifts to a different position with different number and person inflections, but no rule or pattern is suggested, unlike with the *-aj* class. Instead, the unit includes a table (Appendix B) and students are expected to memorize the inflected forms.

In class, the RUSS 001/002 instructor typically refrains from providing too much explicit grammar instruction, opting instead for communicative activities and recasts when students produce ungrammatical utterances. However, when many students seem to be struggling with a particular topic or if they request an explanation, she devotes class time to reviewing the issue explicitly.

Participants

A total of 17 students at UVM enrolled in RUSS 002 in Spring 2020 were the population of this study. All of these students have not been formally instructed in Russian as a second language prior to Fall 2019 in RUSS 001. Additionally, none of them have been to a Russian-speaking country for a period longer than 4 weeks. The course the students are enrolled in is conducted four times per week for an hour each day for a total of 16 weeks. Students are also expected to complete one hour of homework for each class. Therefore, the amount of exposure to Russian input per semester is roughly 128 hours. The course structure and methodology are described in the Flipped Classroom section.

Procedure

There were two types of tests administered to all participants: a Russian proficiency test and a set of five cognitive tests.

In the Russian proficiency test, students were given 40 minutes to complete an assessment that tested their knowledge of both known and novel verbs and their endings. Both sections of the test provided visual stimuli to students that corresponded to each tested verb. Each section had 16 test items, half being known and the other being novel. Both sections mimicked the type frequencies of each verbal class as presented in face-to-face sessions and in the textbook materials (Table 1).

Table 1
Tested Verbal Classes by Section

	-*aj* class	-*i* class	-*a* class
Listening Comprehension (N=16)[4]	11	3	1
Written Production (N=16)	10	5	1

The first portion of the test assessed listening comprehension, asking learners to listen to a simple sentence composed of a personal pronoun and a verb in either the 1st person singular or 3rd person plural and then choose one of two pictures that corresponded to the correct number and person features of the verbal morphology used (Figure 3):

[4] In an attempt to keep the balance of known and novel verbs, an -*e* class verb (*zhit'* "to live") was included in the Listening Comprehension section. The -*e* class is one of the least frequent classes by type frequency, however, *zhit'* is very high in token frequency and students in RUSS 001 and 002 are expected to have mastered its conjugation by the end of the first semester.

1. Профессор: Я завтракаю.
Ja zavtrakaju.
Professor: I am eating breakfast.

А. 

Б. 

Figure 3: Example listening comprehension task

The second portion tested written production. Students read one half of a dialogue and then were asked to complete the dialogue based on the context of the picture stimuli and the previous sentences (Figure 4):

А: Я гуляю. Что они делают?
Ja guljaju. Chto oni delajut?
I am walking. What are they doing?

Б: Они __________.
Oni ___________.
They are ________.

Figure 4: Example written production task

Both listening comprehension and written production modalities were assessed to match the mode of assessment in RUSS 001 and RUSS 002, which includes a high number of visual stimuli.

Five cognitive tests were also administered to be able to correlate students' Russian proficiency with their cognitive capacity, including: working memory (WM), attention, intelligence, and multi-tasking. The cognitive tests were randomly ordered for each participant to reduce any potential ordering effects.

The Automated Operation Span Task (Figure 5) tested individual's WM capacity through several sets composed of simple math equations and letter sequences. First, subjects needed to solve a math equation. In the next screen, a digit is displayed and participants judge the digit as either "true" or "false" based on the answer to the previously solved equation. Immediately after this judgement, a single letter is briefly displayed that the participant needs to commit to memory. This sequence of tasks is repeated for a total of three sets, varying in length from 2 to 7 equation-letter pairs. Lastly, subjects are asked to recall the order of the letter sequences. Feedback is provided to the participants after each set. Both reaction time and correct responses are collected to provide a WM capacity measure.

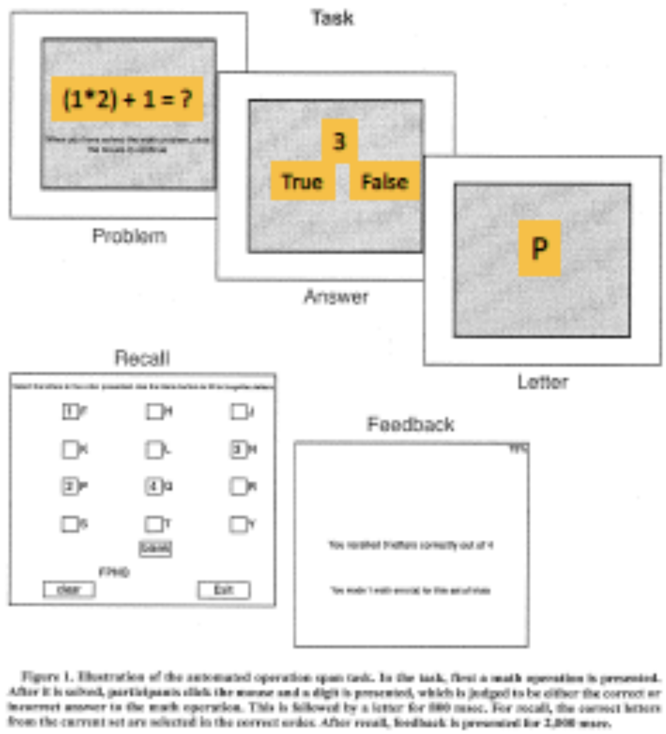


Figure 5: Automated Operation Span Task

The Symmetry Span Task (Figure 6) also assesses WM capacity, albeit through a visual modality. First, subjects experience a distractor task in which they are shown a grid with shaded squares on either side of a central line. Then, participants quickly judge the symmetry of the shape. After this judgment, they are asked to complete the memory task in which they are shown a smaller grid where squares are shaded one by one in a sequence in red. Finally, they are asked to recall the exact order and locations of the shaded squares. Similar to the Operation Span Task, both reaction time and correct responses are collected to generate a WM capacity measure.

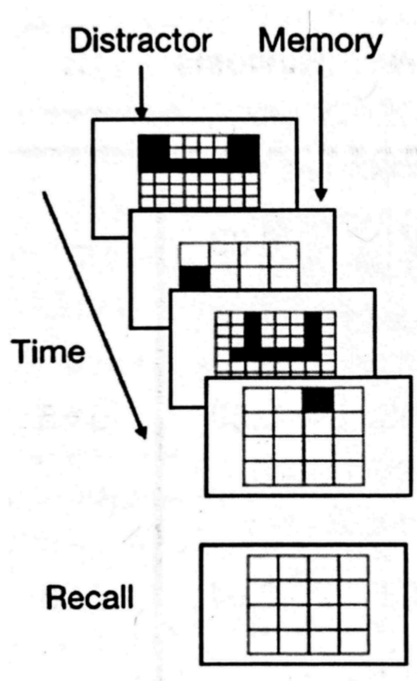


Figure 6: Symmetry Span Task

The Flanker Task (Figure 7) tests individual attention capacity. In this test, participants are briefly shown a line of five arrows, which are either congruent, incongruent, or neutral with respect to the central arrow of the array. In the congruent series, all five arrows point in the same direction as the center arrow. Incongruent series include at least one arrow pointing in the opposite direction as the center arrow. Neutral series include four straight lines surrounding the target arrow. After each series of arrows is displayed, subjects are asked to hit the arrow key corresponding to the direction of the central arrow. This task collects reaction time and correct responses to generate a score called a "Flanker Effect" that reflects attention capacity.

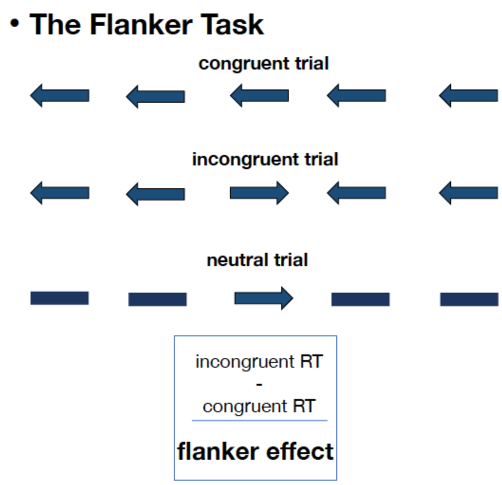


Figure 7: Flanker Task

Raven's Progressive Matrices (Figure 8) tests subjects' fluid, non-verbal intelligence. Eighteen pictures with missing pieces are shown and then participants are asked to select the missing piece from a set of six options. Participants are required to solve all eighteen puzzles in ten minutes. If participants do not complete all puzzles in the timed allotted, they are stopped regardless of their progress. In this task, correct responses are collected to generate a non-verbal intelligence score.

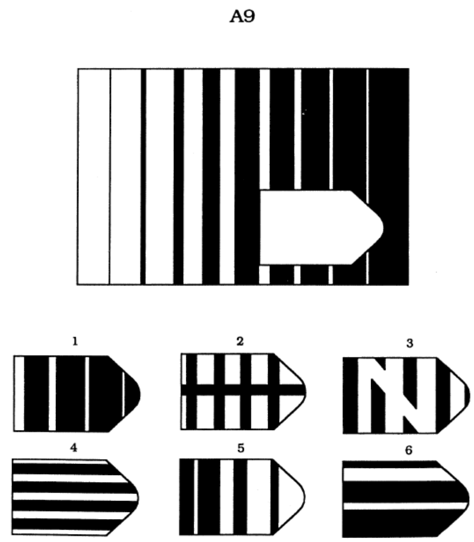


Figure 8: Example Raven's Progressive Matrices test item

Finally, multi-tasking capacity is tested through SynWin (Figure 9), which is an automated multi-tasking test made up of four sub-tasks that the participant is asked to complete simultaneously. In the top left-hand corner, a sequence of six letters is displayed for ten seconds. After the letter sequence disappears, a probe letter appears above. In this sub-task, subjects press the "yes" button if the probe letter is present in the original letter sequence and "no" if it is absent. In the top right-hand corner of the screen, a math equation is displayed. To solve the equation, participants press a plus or minus button below each digit to complete the solution. After finishing the problem, they press the "done" button below the equation. In the lower right-hand section of the screen, subjects are asked to press a red "alert" button every time a high-pitched tone is played. They are explicitly asked to do nothing when the lower tones are played. Finally, in the lower left-hand portion is a "fuel" gauge. Subjects are asked to click on the dial to "refuel" it every time it gets low (they receive more points on their score the longer they wait to refuel without it hitting 0). All of these tasks occur simultaneously and repeat for several cycles during each of the two sets. Participants complete a brief training session and two 5-minute testing sessions. Reaction times and correct responses are collected for all sub-tasks in the two testing sessions to generate a multi-tasking capacity composite measure.

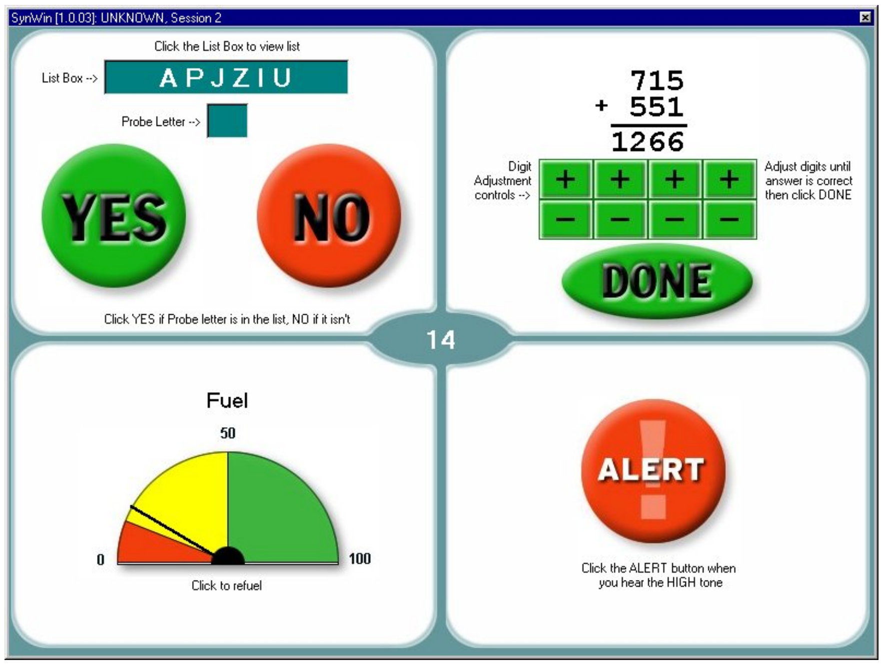


Figure 9: SynWin main interface

Hypotheses

Based on qualitative observations done by the researcher in class, taking into account the overall language performance of the students in RUSS 002, out of the 17 participants, it is expected that around 3 students will perform at high level, and 3 will perform at a low level on the Russian morphology test, while the rest will show average knowledge of Russian verbal morphology. These students' classroom behavior (e.g. amount of participation, general accumulative homework and test scores) suggests who will perform with high and low accuracy on the proficiency test, respectively. In addition, the cognitive profiles of the higher-scoring and lower-scoring participants should justify this variance in proficiency.

Based on findings from Gor & Chernigovskaya (2000-2005) and Tkachenko & Chernigovskaya (2010), it is expected that students will exhibit high accuracy on the *-aj* class. (35) is an example of a prototypical *-aj* class verb in the infinitive. (36) is the first-person singular form of the same verb, which includes the thematic *aj* vowel in its inflectional ending.

(35) *chitat'* "to read"

(36) *ja chitaju* "I read"

We anticipate this finding because *the -aj* class is not only the most frequent verbal class by type frequency, but also because it presents the least amount of morphological complexity (see Russian Verbal Morphology section). Students should demonstrate slightly lower accuracy on *-i* class verbs, which are characterized by their moderate morphological complexity. In (38), we can see that the verb undergoes a consonant mutation from its' infinitive form in (37), just one measure of complexity that can make this paradigm more difficult for students.

(37) *ljubit'* "to love"

(38) *ja ljublju* "I love"

However, we predict that the difference between performance on this class and the *-aj* class may not be statistically significant due to the high frequency of *-i* class verbs in their input. If students demonstrate higher rates of accuracy and generalization of the *-i* class than the more frequent *-aj* class, the thematic vowel *i* will be taken into account (Tkachenko & Chernigovskaya, 2010).

In both *-aj* and *-a* class verbs, the vowel in the infinitive is often *a.* Thus, *-i* class verbs may be more salient in L2 input, especially in an explicit instructed environment, as the vowel *i* present in the infinitive may be seen as a cue to distinguish the stem from the other conjugation paradigms.

Finally, participants are expected to perform at a lower accuracy rate with *-a* class verbs, not only due to the complexity of the conjugation paradigm, but also due to the low type and token frequency of these verbs in the input.

Patterns of generalization of verbal classes to novel verbal stimuli should mirror performance on learned verbs, therefore, *-aj* and *-i* patterns should be applied more often when generalizing, while *the -a* pattern will likely not be extended to other verbs. In fact, it is expected that *-aj* class morphology should be generalized to novel *-a* verbs, due to the low morphological complexity and high type and token frequency of *-aj* class verbs in the input, competing with the complex, low frequency *-a* class.

As for the cognitive variables in this study, it is predicted that WM capacity should account for the highest amount of variance in the overall accuracy on learned verbal stimuli and generalizations to novel verbs. As discussed in the WM section, this capacity includes processing and storage capabilities and therefore should provide learners with the ability to extract patterns from the input to apply to novel utterances. Learners with greater WM capacities can not only hold more information in their short-term store, but also process more information in real-time for eventual conversion into long-term memory. As input filters in through speech in class or through reading text at home, students with greater storage capacities can hold more units of language to be analyzed through their accumulated explicit knowledge. The processing component is equally crucial for analyzing novel verbs, as it allows learners to analyze them and apply previously learned, crystallized explicit rules of Russian conjugation to these new stimuli. More specifically, the effect of WM capacity on performance on morphologically complex and low-frequency verbal classes, particularly the *-a* class, should prove significant. These less frequent, more complex verbal classes impose upon students a need to link specific stems to conjugation paradigms, as Kempe et al. (2010) found for Russian gender morphology, and thus, the construction of these links should be predicted to a reasonable extent by WM capacity. Additionally, because participants are explicitly exposed to grammatical information in this instructional setting, WM should be highly predictive of accuracy, as students are told, through homework assignments and in class, to pay attention to specific elements of the input, which stimulates the development of grammatical schema[5] (see McDonough & Trofimovich, 2016).

Recent studies have shown that WM is a significant predictor of multitasking abilities (Hambrick, Oswald, Darowski, Rench & Brou, 2010; Redick, Shipstead, Meier, Montroy, Hicks, Unsworth, Kane, Hambrick & Engle, 2016). Therefore, it is also expected that along with WM capacity, multitasking ability should account for a reasonable amount of variance in the morphology test results based on its connection to WM capacity and the demands of a communicative FC environment during f2f weekly sessions.

Similarly, fluid intelligence has been found to be a significant predictor of multitasking ability (r=.76), although the strength of this relationship is slightly weaker than that of WM and

[5] The instructions on the proficiency test were also designed to elicit this explicit information.

multitasking (r=.77) (Redick et al, 2016). Thus, it is expected that fluid intelligence should account for some amount of variance in the verbal morphology test scores.

Finally, we anticipate that attention capacity should account for some variance in the verbal morphology test scores, as many researchers consider attention to be one and the same with the executive control function of WM (see discussion on Kapa & Colombo, 2014 in the WM section). However, it is important to note that in Redick et al's (2016) study, attention accounted for the least amount of variance in multitasking capacity, which is heavily implicated in WM capacity.

Results

Russian Verbal Morphology Test

Comprehension Task

On the comprehension half of the test, all but one student performed at 100% accuracy (at ceiling). While this may seem to indicate that most students have a high level of mastery on all three tested verbal classes, there are several variables to consider in the interpretation of these results, which are considered in the Discussion section.

Written Production Task

In contrast to the comprehension task, learners exhibited much more variance in their written production scores (see Appendix C for raw scores for each participant). Items were counted as correct if the student satisfied the following criteria:

(35) Generalized the verb to the correct verbal class

(36) Produced the correct verbal form (person and number)

(37) Wrote the verb correctly with minimal orthographic errors

Orthographic errors that were considered insignificant were ones that retained the thematic vowel of the verbal class' conjugation paradigm, contained very minimal errors (e.g. writing *a* for *o*, as unstressed *o* can be pronounced as /a/ or /ə/), and did not confuse person or number. For example, **ja zakazyvauju* ("I order") was counted as correct as the learner retained the thematic *aj* vowel, albeit with the addition of the extra vowel *u.* However, the same learner produced **oni sprashivujut* ("They are asking"). This was counted as incorrect because of the complete lack of the thematic *aj* vowel even though the -*aj* class 3rd person plural ending -*jut* is intact.

No learner reached 100% accuracy on this portion of the test, although 2 students reached 93.7% accuracy (15 out 16 correct). There was a total of 2 students who performed at 50% accuracy or lower, roughly aligning with the predictions regarding lower-level performers in the Russian classroom (see Hypotheses). 10 students, or 58.8% of the participants, performed at 75% accuracy or higher. A total of 7 students (41.2% of the population) answered 14 out of 16 or more items correctly, performing at 87.5% accuracy or higher. Tables 2 and 3 summarize this data.

Table 2
Accuracy Scores on the Written Production Task

Number of students	93.7% accuracy	87.5%+ accuracy	75%+ accuracy	50% or lower accuracy
Total (N=17)	2	7	10	2

Table 3
Descriptive Statistics: Written Production Scores

	Mean	% Accuracy	Standard Deviation
Written Production	11.8	73.7%	2.8

As both components of the task were designed to map onto the type frequencies of verbal classes present in the input, only one -*a* class verb was tested per component. A total of 9 students (53%) were able to generalize the -*a* class pattern to the single -*a* class verb on the Written Production Task.

Cognitive Testing

Before discussing the significant results found among all cognitive and morphological variables, it is important to note that a total of 4 participants (3001, 3005, 3008, and 3016) were excluded from the ultimate analysis. All of these participants demonstrated outlier behavior on one or more of the cognitive tests, which significantly impacted the resulting correlations. Participants 3005, 3008, and 3016, for example, all scored significantly lower than the rest of the group on the SynWin. Sparing these participants, the average score on the SynWin for both sessions was 707.8. This is ultimately why they were eliminated from the dataset for further analysis so as not to include outliers that negatively affected resulting correlations. Participant 3001 was removed for similar reasons regarding outlier behavior, particularly due to their absolute scores on the Automated Operation Span Task and Symmetry Span, which were significantly lower than other participants' scores, at 16 and 2, respectively.

After removing these 4 outliers, the scores on all tasks were subjected to Pearson's test of correlation. There was a significant correlation of (r=.61) between average scores on the

Operation Span Task and the scores on the total score for the Russian Morphology Test. While the absolute scores on the Morphology Test reached significance with this measure of WM, the Written Production subsection of the Morphology Test approached significance (r=.54) with the Operation Span Task.

Likewise, the Flanker Effect, or the difference between participants' scores on the incongruent and congruent trials of the Flanker Task, was significantly *negatively* correlated with both the absolute scores on the Morphology Test (r=-.59) and the Written Production subsection (r=-.59). Therefore, along with a significant relationship with a more "traditional" WM task (the Operation Span), we can observe a significant relationship between attention (or the inhibitory control executive function) and L2 knowledge of verbal morphology.

The scores on the Raven's Progressive Matrices were significantly *negatively* correlated (r=-.59) with participants' reaction times on the Symmetry Span Task, meaning that participants with faster reaction times on this specific test of WM scored higher on our fluid intelligence test.

Both the individual first and second sessions of the SynWin approached a significant relationship (r=.52 and r=.54, respectively) with scores on Raven's Progressive Matrices. The average scores on both sessions of the SynWin, however *did* reach significance in correlation with our fluid intelligence test, with a correlation of r=.56.

The absolute scores on the Symmetry Span Task, one of our WM measures, reached a highly significant correlation with the first session of the SynWin at r=.70. The second session was also significant, with a correlation of r=.61. The average scores between both sessions of the SynWin also reached significance with the Symmetry Span Task with a correlation of r=.69.

Similar to the relationships among the absolute scores on the Symmetry Span, the partial scores on the same task also reached significance for the individual sessions of the SynWin test (r=.78 and r=.61, respectively) and the average scores between both sessions (r=.74).

Tables 4-6 show correlations that are significant or approach significance.

Table 4
Correlations with the Russian Verbal Morphology Test

	AOSpan Partial	Flanker Effect
Written Production	0.538 0.058 13	r=-0.589* 0.034 13
MorphTotal	r=0.610* 0.027 13	r=-0.590* 0.034 13

Table 5
Correlations Between the Symmetry Span and SynWin

	SynWin 1	SynWin 2	SynWin Av.
SSpan Absolute	r=0.705** 0.007 13	r=0.612* 0.026 13	r=0.697** 0.008 13
SSpan Partial	r=0.782** 0.002 13	r=0.617* 0.025 13	r=0.740** 0.004 13

Table 6
Correlations Between Raven's Progressive Matrices and SynWin

	SynWin 1	SynWin 2	SynWin Av.
Raven's	r=0.521 0.068 13	r=0.539 0.057 13	r=0.561* 0.046 13

Discussion

Firstly, a discussion on the results from the Russian Verbal Morphology test is in order. On the Comprehension Task, all but one student performed at 100% accuracy. This figure looks surprising on the surface, but in reality there are several variables present in the students' classroom that may explain this result. the structure of the students' classroom places equal importance on listening comprehension and written production capabilities. During class and on examinations, students are expected to answer comprehension questions in English and in Russian in addition to completing dictations or writing down what the instructor says. At the same time, they are also tested on their ability to produce sentences from prompts and occasionally, uninflected word banks. Therefore, both comprehension and production abilities must be analyzed together to determine true proficiency.

Secondly and most importantly, the listening prompts contained redundant information. In Russian, unlike some languages (e.g. Spanish), the inflected verb must be accompanied by its pronoun. When students heard these prompts, therefore, they heard the more salient word, the pronoun, followed by a redundant verb form containing the proper inflection for person and number. As the default *-aj* class endings for 1st person singular "I" and "they" 3rd plural are distinguished by just a single phoneme (*-ju* versus *-jut*), students likely relied on the more salient pronoun distinction to determine the correct answer for each prompt (*ja* versus *oni*) in the Comprehension section of the Verbal Morphology Test.

As for the Written Production Task, the fact that so many students were able to generalize the low frequency *-a* class pattern correctly may seem surprising. However, as suggested in one of the hypotheses, it is reasonable to assume that learners have internalized the *-a* class conjugation paradigm as explicit knowledge because it is emphasized in the textbook as an "irregular" pattern that needs to be committed to memory. This is similar to Tkachenko and Chernigovskaya's (2010) findings, in that the *-a* class proved more salient due to the explicit emphasis on its "irregularity", despite the fact that it is a very low type frequency verb. Additionally, the verb in question is provided in the 1st person singular, from which students would only need to apply a single letter (*-t*), to produce the correct form, *oni plachut* ("they cry"). Thus, it may be the case that the 1st person singular served as a primer for the 3rd person plural, as it often does for *-aj* forms that behave in the same way (*ja dumaju* → *oni dumajut*, "I think" → "They think").

However, the fact that students were able to generalize *-a* to other verbs (albeit incorrectly) needs to be accounted for. This was particularly prevalent in Item 4 of the Written Production Task, which prompted students with the 1st person singular form of "to fly," *ja lechu*. Although this form has many of the hallmarks of an *-a* class verb (consonant mutation, stress shift), it in fact belongs to the *-i* class, which contains many of the same features. The difference between the two, however, is the continuation of the consonant mutation from the 1st person singular form through to the 3rd person plural. For example, the verbs *letet'* "to fly" and *plakat'* "to cry":

(39) *ja lechu*

(40) *oni letjat*

(41) *ja plachu*

(42) *oni plachut*

In (40), we can observe that the *t > ch* mutation in this *-i* class verb is not preserved in the 3rd person plural form, while in the *-a* class verb in (42), the consonant mutation is preserved through the 3rd person plural. In both verbs, however, the 1st person singular reflects the consonant mutation and stress shift indicative of both verbal classes. When these forms are presented as stimuli instead of the more transparent infinitives or past tense forms[6], learners are forced to make a choice between two very similar conjugation paradigms when presented with the 1st person singular. It could be the case that the 9 students who generalized the *-a* class paradigm to *letet'* "to fly", chose to do so for the same reason why they were so successful in producing the *-a* class correctly for *plakat'* "to cry". The explicit knowledge of this seemingly "irregular" pattern, its endings, and consonant mutations may signal to them that this is the right choice. Therefore, rather than relying on frequency, learners default to their declarative knowledge about Russian conjugation.

Additionally, while it cannot be claimed that these 9 learners who generalized *-a* (incorrectly) to the *-i* class verb *letet'* "to fly" have "mastered" the paradigm, it certainly indicates that they possess explicit knowledge pertaining to how the verbal class is conjugated.

---

[6] The vowel in the stem in both the past tense and infinitive forms of both *-i* and *-a* class verbs is transparent. For example: *ja letel* "I flew" and *ja plakal* "I cried". While it is reasonable for *plakal* to be generalized as an *-aj* class verb in the present tense, due to the stem *-a* vowel it shares in common with the *-aj* class, the *-e* in *letel* is a strong cue to learners that it is *not* a *-a* or *-aj* class verb, as that stem vowel is never found in those respective conjugation paradigms.

Otherwise, all learners would either generalize the verb correctly as the -*i* class[7] or incorrectly as the -*aj* class[8]. No participants generalized -*aj* for this verb.

Turning now to the relationships between the language proficiency results and cognitive data, let us first consider that the partial scores on the Operation Span Task correlated with the overall scores on the Morphology Test (r=.54). As we observed in the results, the Oral Comprehension Task resulted in very little variance. As such, it is important to note, that this correlation is driven by the Written Production scores, which varied among the population to a much greater extent than that of the Oral Comprehension scores. The scores on the Written Production Task ultimately approached significance with the Operation Span Task. From these results, we can conclude that the hypothesis on the connection between WM capacity and Russian proficiency is adequately supported by this data. To explain the significance of this result, however, let us consider Kempe, et al.'s (2010) study, in which L2 Russian learners with higher WM capacities were more successful in recognizing grammatical gender *and* generalizing gender categories to novel items. Kempe and colleagues asserted that, because these learners possessed greater WM capacities, they were more successful in extracting patterns from their L2 input and later applying them to novel items. As discussed in the WM section, it is understood that this capacity is limited, and thus individuals with larger capacities are able to hold more incoming information within their store, which can later be analyzed for rule generation. It is likely that participants in this study who scored high on WM measures were likewise able to hold more L2 input in their WM store, allowing for them to more successfully

[7] The -*i* class commonly features infinitives with -*e* as the stem vowel, such as *smotret'* "to watch", *videt'* "to see", and more. It can also contain -*i* as the stem vowel: *govorit'* "to talk", *ljubit'* "to love".

extract patterns for later production, particularly given the written modality of this section of the test that, although timed, allows for the deployment of explicit knowledge and analogical reasoning.

The attention measure (Flanker Effect) and the global scores on the Morphology Test exhibited a significant relationship (r=-.58), suggesting that learners with greater attention capacity are more successful in their production and comprehension of Russian verbal morphology. Kapa & Colombo's (2014) study, found that L2 learners with higher levels of attentional control were more successful in acquiring L2 morphosyntax. To recall our discussion on WM, this result is not unexpected, as many consider attentional control to be one and the same with the central executive, or the central module in the WM model (Baddeley & Hitch, 1974). This also follows Linck et al.'s (2014) findings from a meta-analysis of WM in L2 learning research, in that measures connected to the central executive are the best predictors of L2 performance. Perhaps our participants with higher attentional control, measured by the Flanker Task, were able to better inhibit distractors (i.e. the "inner voice" of English translation) while processing L2 Russian input.

Significant correlations were also found among the cognitive measures themselves: between the Symmetry Span Task (WM) and SynWin (multi-tasking) (r=.69), and Raven's Progressive Matrices (intelligence) and SynWin (r=.56), indicating that all measures of cognitive capacity intercorrelate to different degrees. The significant relationship between the Symmetry Span and SynWin suggests that these learners' WM capacity, specifically their visuospatial capabilities, is connected to their multi-tasking capacity, which in itself is considered to be a measure of executive control (Redick et al., 2016). Although these measures fail to reach significance with our measures of proficiency, it is important to acknowledge that these correlations among these cognitive variables are observed in previous studies (Redick et al., 2016), indicating the validity of the tests used in this thesis.

The Symmetry Span Task and the Morphology Test scores did not reach a significant correlation (r=.01). Recall that there was, however, a significant relationship between the Operation Span Task, a different measure of WM capacity, and the test scores. This discrepancy could be accounted for the fact that the Operation Span Task uses letters of the alphabet in its test, although it strives to be language independent. The truly language-independent task, the Symmetry Span, has no relationship with our proficiency measure, which may provide evidence to support the claim that there is a connection between type of task and ability to handle L2 input.

In the primary analysis of the results, the lack of a significant relationship between fluid intelligence (Raven's Progressive Matrices) and the Russian Verbal Morphology Test (r=.21) provides evidence *against* one of the hypotheses posed in this study. Before examining this finding further, it is important to note that while Redick et al.'s (2016) study on multi-tasking indicated that fluid intelligence is a significant predictor of multi-tasking capacity (r=.76), the study also found that multi-tasking was well mediated by WM (r=.77) and attention (r=.61) scores. However, in previous studies on the acquisition Russian grammar, higher intelligence scores have predicted performance on novel items specifically. Kempe et al. (2010) propose that these learners are able to more successfully extract patterns from the input they are afforded, and thus are able to successfully generalize these patterns to novel items. Taking these findings into account, we decided to run a correlational analysis on known and novel verbs tested in the Written Production portion of the test. If what Kempe et al. (2010) found extends to our results, students should exhibit more variance on their novel verb test scores. The results are presented below in Tables 7 and 8.

Table 7
Descriptive Statistics: Known vs. Novel Verbs

	Known	Novel
Mean	6.30	5.76
Standard Deviation	1.54	1.48

Table 8
Correlations Between Raven's Progressive Matrices, the Flanker Effect, & Known vs. Novel Verbs

	Raven's	Flanker Effect
Known Verbs	0.174 0.569 13	0.165 0.590 13
Novel Verbs	0.100 0.744 13	r= -0.813* 0.001 13

Despite distinguishing between the learned status of verbs among the student population, we still did not find a significant relationship between any set of verbs and our fluid intelligence measure. However, upon running these statistics, we did find a highly significant correlation between novel verbs and the Flanker Effect (r=-.81).

Fluid intelligence is the ability to solve new problems without relying on accumulated, explicit knowledge. Because these stimuli require learners to solve a new problem, the observation of a null relationship between the novel verbs and intelligence is thus, quite surprising. However, the variation in scores between known and novel verbs assessed through a paired-samples t-test proved to be not significant (see Tables 9 and 10).

Table 9
Paired Samples T-Test

	Correlation	Sig.
Known Verbs & Novel Verbs	0.07	0.82

Table 10
Paired Samples T-Test: Paired Differences

	Mean	St. Deviation	Std. Error Mean
Known Verbs & Novel Verbs	0.53	2.06	0.57
	95% Confidence Interval of the Difference		
	Lower	Upper	
	-0.71	1.78	
	t	df	Sig. (2-tailed)
	0.94	12	0.36

To determine whether the known or the novel verbs drove the relationship between the Morphology Test and the cognitive measures, we ran additional correlations (Table 11). The novel verbs seem to be the motivator behind the correlations between scores on the Morphology Test and the cognitive measures. For example, the relationship between the Flanker Task and scores on the Written Production Task only reached significance for the novel items (r=-.81) and not for the known items . Therefore, how well or poorly students perform *only* on novel items determined the significant relationship with our attention measure. More specifically, it is the incongruent trials on the Flanker Task that drive the correlations between Morphology scores and this attention measure (r=-.62). It is also important to note that the absolute scores on the Morphology Test correlate only with the novel verbs (r=.79).

Table 11
Correlations Between Known & Novel Verbs and the Morphology Test

	Novel Verbs	MorphTotal	Flanker Incongruent	Flanker Effect
Written Production	r=-0.816** 0.001 13			
Known Verbs		0.268 0.377 13		0.165 0.590 13
Novel Verbs		r=0.797** 0.001 13	r=-0.532 0.062 13	r=-0.813** 0.001 13
Flanker Incongruent				r=-0.627* 0.022 13

Looking more closely at the makeup of the proficiency test, heavy weight (10 out 16 total items) was given to the *-aj* class, the most frequent verb class by type frequency in the RUSS 001 and 002 classroom, and even in the Russian language as a whole. Only 5 *-i* class verbs, 3 of which were known and 2 of which were novel, and only one *-a* class verb (novel) were tested, meaning that scores overall for both the known and novel categories are heavily biased toward the most productive class. Although there is considerable variance in WM capacity in our population, performance on these *-aj* class verbs is consistent. Table 12 shows that students performed at 89.2% accuracy on both the known *and* novel *-aj* class items, suggesting that most students have developed the explicit rules necessary to generalize this pattern to novel stimuli.

Table 12
Descriptive Statistics: -aj *Class Verbs on the Written Production Task*

	Mean	Standard Deviation	Average Accuracy (% correct)
All *-aj* class verbs N=10	8.92	1.25	89.2%
Known *-aj* class verbs N=5	4.46	0.66	89.2%
Novel *-aj* class verbs N=5	4.46	0.87	89.2%

As for performance on the *-i* class items, the results revealed a different story (see Table 13). Accuracy on the known verbs of this class reached 61.3% percent, while accuracy on the novel verbs only measured at 38%. It is not only clear that there is a with *-i* class verbs, but also that overall accuracy (52.2%) is much lower than that of the *-aj* class.

Table 13
Descriptive Statistics: -i *Class Verbs on the Written Production Task*

	Mean	Standard Deviation	Average Accuracy (% correct)
All *-i* class verbs N=5	2.61	1.44	52.2%
Known *-i* class verbs N=3	1.84	1.21	61.3%
Novel *-i* class verbs N=2	0.76	0.87	38%

Now, let us consider the relationship between novel verb scores and the Flanker Effect. First, students with greater attention capacities are able to better filter out distractors and devote the focus of attention (Cowan, 2005) on incoming input. As this input is held in the WM store, it can be processed and analyzed. During these processes, students may search for commonalities among learned verbs and these novel stimuli, drawing from previously learned explicit rules. Students with greater WM and attention capacities are typically more successful in these endeavors, which allows for them to perform at higher accuracy on novel stimuli. Unlike novel verbs, known verbs are likely stored in long-term memory, where students can retrieve those units of language as a chunk. This may explain why we observe no significant relationship between attention and known verbs, as WM plays a lesser role in the retrieval of learned("crystalized") units of language from long-term memory.

Secondly, let us consider the effects of students' exposure sessions in RUSS 001 and 002. In the grammar explanations and occasionally in class, students are provided with explicit information regarding grammatical concepts. Additionally, in the instructions, and verbally before administering the Written Production section of our morphology test, we told students to pay attention to verb endings. McDonough & Trofimovich (2016) claim that WM effects are only observed when learners are told explicitly to pay attention to elements in the input. Telling a student to use their explicit knowledge in cases such as these raises students' level of attention, particularly when it comes to verbs they have not previously encountered. When faced with such stimuli, they must deploy the executive control component of their WM (aka attention) to notice any clues in the model provided, in our study, either the 3rd person plural or 1st person singular. This may explain why we see students with high scores on the Flanker Task performing more accurately on the Written Production section.

As for the role of teaching methodology in student success regarding Russian verbal morphology, it is safe to conclude that there is little to no effect. In Li et al.'s (2019) study, students were exposed to different types of instruction to assess whether WM effects would be generated by different types of intervention. Students who experienced only pre-task explicit grammar instruction were observed to have no WM effects on performance. If this is the case, we should have seen little to no WM effects among *our* student population. In reality, while RUSS 001 and 002 *are* flipped classrooms (FCs), in which explicit grammar instruction is self-directed by students at home *prior* to communicative practice, explicit instruction is still delivered in class. The majority of class time is devoted to encouraging students to use their Russian language skills, but occasionally it is fundamentally necessary to intervene and explain grammatical topics when students consistently produce errors or experience inhibition while producing the language. Instructors view this strategy not as a departure from the fundamentals of the communicative method, but more so as a way to blend instruction types to ensure global understanding among the student population. In language classrooms, students regularly ask questions that target explicit explanations in class (e.g. why does the *s* in *pisat'* change to *sh* in *pishu*). As an instructor, one should not deny students of the answers they are seeking to find. If some of these answers target information meant for homework, one cannot simply ignore the question and move on to more practice, as it is necessary for students to understand what they produce. In fact, delivering feedback to the entire class assures that all students have the opportunity to understand targeted forms.

Conclusion

Understanding now that the FC studied in this thesis occasionally incorporates explicit instruction in class, one should be cautious when comparing our results to those in Li et al.'s (2019) study, as our learning environment operationalization is not strictly controlled. In Li's own study[9], however, for the condition in which both pre-task instruction and in-task feedback were delivered, WM effects *were* found, similar to our own findings in this thesis. In that study, the researchers suggest that the learners under the pre- and in-task instruction condition were required to take in language input and corrective feedback simultaneously (Li et al., 2019), which demands much more WM resources versus the group that did not experience instruction in-task. Our learners, by extension, seem to be impacted by WM capacity in the same way, as they experience similar learning conditions in the RUSS 001 and 002 classroom.

[9] This study was conducted in a laboratory environment.

Since we have determined that students with higher WM and attention capacities perform better in this classroom, now the matter of how to accommodate those with different cognitive capacities needs to be addressed.

Students with lower WM and attention capacities are typically unable to hold as much information presented in the L2 or retain as much focus in-task, when compared to their higher-capacity peers. Although these students have the same exposure to explicit instruction as their counterparts, it may be that they are not able to access the required knowledge from their long-term memory store as efficiently due to their lower WM capacity. Offering supplemental instruction or tutoring beyond the classroom may help these students who seem to struggle. One-on-one interaction in which the student is able to request specific types of explanation and practice may help bootstrap the instruction they receive in class and at home. This type of opportunity may also break down the walls of inhibition for some of these lower-capacity students who feel uncomfortable participating in class. Supplemental settings in which the student is only faced with the instructor or a Teaching Assistant has the potential to mitigate these affective variables in addition to scaffolding explicit knowledge needed to produce Russian verbal morphology, and more broadly, Russian grammar as a whole. Simply put, these students may require more time to take in L2 Russian input and derive the same explicit knowledge as their peers.

Despite the fact that our research confirmed some of our hypotheses, there were a few limitations present in the study. In terms of the teaching methodology at the center of this thesis, it is important to note that there could be no control over how much time students spent on reading explicit grammar instruction and doing the corresponding activities at home.

Therefore, it is impossible for us to determine how much time the average student spent on memorizing forms and practicing patterns on their own. The online textbook has no function to check how much time students spend on the grammar readings and activities. Secondly, the design of the Listening Comprehension section of the Russian Verbal Morphology Test included redundant test items. Instead of primarily attending to word-final inflected morphemes, the design allowed for students to focus more on the personal pronoun provided. Since Russian, unlike Romance languages such as Spanish, requires a noun or personal pronoun before an inflected verb, the verbal ending provides redundant information. This very fact is likely the reason why we found students performing at ceiling on this section (spare one student). However, although the design was flawed, we saw no other way to test oral comprehension, as including just the verb in isolation would have provided students with ungrammatical sentences.

Suggestions for Future Research

For future research specific to this grammatical feature of Russian, we would suggest testing students after the end of their second semester. At this point, students will have learned verbs from all four of the verbal classes tested by Tkachenko & Chernigovskaya (2010), including the *-ova* class, which has a complex conjugation paradigm along with high type frequency. Being able to test this specific verbal class would provide more insight into student performance, as well as generating more material to later be analyzed in conjunction with measures of individual differences. As for investigating the role of flipped classroom teaching methodology, it would be worth implementing experiments in a classroom in which students are required to spend a certain amount of time on explicit grammar work at home and where a tracking system can quantify students' time on-task for the research purposes. This condition would allow for more control over extraneous variables present in the current study and provide for more conclusive findings as to the benefits and drawbacks of the FC methodology.

References

Baddeley, A. (1981). The concept of working memory: A view of its current state and probable future development. *Cognition, 10*(1-3), 17-23.

Baddeley, A. (2000). The episodic buffer: A new component of working memory?. *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A. (2007). Introduction and overview. In A. Baddeley (Eds.), *Working memory, thought, and action.* Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol 8, pp. 47-89). Academic Press.

Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition. *An Interdisciplinary Journal of the Language Sciences, 47*(2), 383-411.

Chernigovskaya, T., & Gor, K. (2000). The complexity of paradigm and input frequencies in native and second language verbal processing: Evidence from Russian. *Language and Speech Behavior, 3*(II), 20-37.

Correa, M. (2015). Flipping the foreign language classroom and critical pedagogies: A (new) old trend. *Higher Education for the Future, 2*(2), 114-125.

DeBenedette, L., Comer, W. J., Smyslova, A., & Perkins, J. (2019). *Mezhdu Nami.*

Ellis, N. C. (1998). Emergentism, connectionism, and language learning. *Language Learning, 48*(4), 631-664.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143-188.

Ellis, N. C. (2008). The dynamics of second language emergence: Cycle of language use, language change, and language acquisition. *The Modern Language Journal, 92*(2), 232-249.

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer-distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition, 19*, 145-171.

Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philip & H. Reinder (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-25). Multilingual Matters.

Gor, K., & Chernigovskaya, T. (2001). Rules in processing of Russian verbal morphology. In G. Zybatow, U. Junghanns, G. Mehlhorn, & L. Szucsich (Eds.), *Current Issues in Formal Slavic Linguistics* (pp. 528-535). Peter Lang.

Gor, K., & Chernigovskaya, T. (2003). Mental lexicon structure in L1 and L2 acquisition: Russian evidence. *Glossos,* (4), 1-31.

Gor, K., & Chernigovskaya, T. (2005). Formal instruction and the acquisition of verbal morphology. In A. Housen & M. Pierrard (Eds.), *Investigation in instructed second language acquisition* (pp. 131-164). Mouton de Gruyter.

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology,* 24, 1149-1167.

Jakobson, R. (1948). Russian conjugation. *WORD, 4*(3), 155-167.

Kapa, L., & Colombo, J. (2014). Executive function predicts artificial language learning. *Journal of Memory and Language, 76,* 237-252.

Kempe, V., & Brooks, P. J. (2005). The role of diminutives in the acquisition of Russian gender: Can elements of child-directed speech aid in learning morphology? *Language Learning,* 55, 139-176.

Kempe, V., & Brooks, P. J. (2008). Second language learning of complex inflectional systems. *Language Learning, 58*, 703-746.

Kempe, V., Brooks, P. J., & Kharkhurin, A. (201). Cognitive predictors of generalization of Russian grammatical gender categories. *Language Learning, 60*(1), 127-153.

Langacker, R. W. (1987). *Foundations of cognitive grammar, vol. 1: Theoretical prerequisites.* Stanford University Press.

Li, S. (2017). Cognitive differences and ISLA. In S. Loewen & M. Sato, The Routledge handbook of instructed second language acquisition (pp. 396-417). Routledge.

Li, S., Ellis, R., & Zhu, Y. (2019). The associations between cognitive ability and L2 development under five different instructional conditions. *Applied Psycholinguistics,* 40, 693-722.

Linck, J., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*, 861-883.

Linck, J. A., & Weiss, D. J. (2011). Working memory predicts the acquisition of explicit L2 knowledge In C. Sanz & R. P. Leow, *Implicit and Explicit Language Learning* (pp. 101-113). Georgetown University Press.

Linck, J. A., & Weiss, D. J. (2015). Can working memory and inhibitory control predict second language learning in the classroom? *SAGE Open, 5*(4), 1-11.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). John Benjamins.

Martin, K., & Ellis, N. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*, 379-413.

Matcovich, P. (2001). Regular inflection in the mental lexicon: Evidence from Italian In H. G. Simonsen & R. T. Endresen (Eds.), *A cognitive approach to the verb: Morphological and constructional perspectives.* Mouton de Gruyter.

McDonough, K., & Trofimovich, P. (2016). The role of statistical learning and working memory in L2 speakers' pattern learning. *The Modern Language Journal, 100*(2), 428-445.

McWhinney, B. (2013). Competition model. In P. Robinson (Ed.), *The Routledge Encyclopedia of Second Language Acquisition* (pp. 100-103). Routledge.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude In A. F. Healy & L. E. Bourne (Eds.), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention* (pp. 339-364). Lawrence Erlbaum Associates.

Pawlak, M., & Biedrón, A. (2019). Verbal working memory as a predictor of explicit and implicit knowledge of English passive voice. *Journal of Second Language Studies, 2*(2), 258-280.

Prefume, Y. E. (2015). *Exploring a flipped language classroom approach in a Japanese language classroom: A mixed methods study.* [Doctoral dissertation, Baylor University]. Baylor University.

Ragnarsdóttir, H., Simonsen, H. G., & Plunkett, K. (1999). The acquisition of past tense morphology in Icelandic and Norwegian children: An experimental study. *Journal of Child Language, 26*(3), 577-618.

Redick, T. S., Shipstead, Z., Meier, M. E., Montroy, J. J., Hicks, K. L., Unsworths, N., Kane, M. J., Hambrick, D. Z., & Engle, R. W. (2016). Cognitive predictors of a common multitasking ability: Contributions from working memory, attention control, and intelligence. *Journal of Experimental Psychology: General, 145*(11), 1473-1492.

Roncaglia-Denissen, M. P., & Kotz, S. A. (2016). What does neuroimaging tell us about morphosyntactic processing in the brain of second language learners? *Bilingualism: Language and Cognition, 19*(4), 665-673.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua, 120*, 2022-2039.

Sanz, C., Lin, H., Lado, B., Stafford, C. A., & Bowden, H. W. (2014). One size fits all? Learning conditions and working memory capacity in ab initio language development. *Applied Linguistics,* 1-26.

Tagarelli, K. M., Borges-Mota, M., Rebuschat, P. (2011). The role of working memory in implicit and explicit language learning, *Proceedings of the Annual Meeting of the Cognitive Science Society, 33*(33), 2061-2066.

Tkachenko, E., & Chernigovskaya, T. (2010). Input frequencies in processing of verbal morphology in L1 and L2: Evidence from Russian. *Russian in Contrast. Lexicon, 2*(2), 281-318.

Tonkin, K., Page, S., & Forsey, M. (2019). Managing cognitive load with a flipped language class: An ethnographic study of the student experience. *Foreign Language Annals, 52*(3), 551-575.

Tyler, A. (2010). Usage-based approaches to language and their applications to second language learning. *Annual Review of Applied Linguistics, 30*, 270-291.

Weissheimer, J., & Mota, M. B. (2009). Individual differences in working memory capacity and the development of L2 speech production. *Issues in Applied Linguistics, 17*(2), 93-112.

Wen, E. (2016). *Working memory and second language learning: Towards an integrated approach.* Multilingual Matters.

Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: building and supporting a case for "working memory as language aptitude". *Ilha Do Desterro: A Journal of English Language, Literature in English and Cultural Studies, 60*, 15-43.

Appendices

Appendix A: Gor & Chernigovskaya Studies

Table 14
Chronology of Research Questions

	RQ #1	RQ #2	RQ #3	RQ#4
2001	What is the default pattern? Which conjugational patterns are more likely to be generalized to other verb classes?	Are generalizations influenced by type frequencies or complexity of the verbal class?	What is the role of morphological cues in verbal processing?	Are the rules applied in a set, or can they be disassociated?
2003	Is there a developmental tendency in child L1 acquisition of this feature?	Does processing in L2 learners match the processing of any of the child L1 age groups?	Which population relies more on associative patterning?	
2004	Is there a developmental tendency in child L1 acquisition of this feature?	Does processing in L2 learners match the processing of any of the child L1 age groups?	Which population relies more on associative patterning?	
2005	Does explicit instruction result in successful learning of these rules (reflected on written tests)?	Does explicit instruction facilitate the development of native-like processing strategies in L2 learners?	What is the role the role of input frequencies in L2 processing?	

Tkachenko & Chernigov-skaya, 2010	Classes with high type frequency should not exhibit any differences between child L1 and L2 learners.	The most frequent patterns should be applied more frequently to nonce verbs and serve as the basis of overgeneral-ization.	The class with lowest type frequency should be acquired later by child L1 speakers and L2 learners may have difficulty acquiring this pattern.	There should be differences in the acquisition of *-ova* because the type frequencies are different in L1 and L2 environments.

Table 15
Chronology of Experimental Conditions

	L1 Group	L2 Group	# of Verbal Classes	# of Verbs	Stimulus Form	Mode
2000	27 adults	15 students	9	L1: 48 nonce L2: 48 real	Past tense plural	Oral
2001	27 adults	n/a	9	L1: 48 nonce	Past tense Plural	Oral
2003	5 4 y.o. 9 5 y.o. 6 6 y.o.	20 students	4	L1: 40 real, 20 nonce L2: 40 real, 20 nonce	Past tense plural & infinitive	Oral
2004	5 4 y.o. 9 5 y.o. 6 6 y.o.	20 students	4	L1: 40 real, 20 nonce L2: 40 real, 20 nonce	Past tense plural & infinitive	Oral
2005	27 adults	15 students	9	L1: 48 nonce L2: 46 real (written) 48 real (oral)	Basic stem, past tense plural, & 3rd person plural	L1: Oral L2: Writt-en & Oral
Tkachenko & Chernigo-vskaya, 2010	30 4 y.o. 30 6 y.o. 21 8 y.o.	21 L1 Norwegian speakers	4	L1: 40 real, 40 nonce L2: 40 real, 40 nonce	Past tense plural & infinitive	Oral

Appendix B: RUSS 001/002 and Mezhdu Nami

Mezhdu Nami Homework

А что де́лают их хозя́ева?

Зо́я Степа́новна де́лает сала́ты.

Ри́мма Ю́рьевна до́ма и убира́ет кварти́ру.

Мара́т Аза́тович рабо́тает.

А где Светла́на Бори́совна? Что она́ сейча́с де́лает? Мы не зна́ем. Э́то большо́й секре́т.

Nemnogo o jazyke: Default -*aj* class conjugation table

The vast majority of first conjugation verbs have stems like the one for "to read," whose forms you see on the right. These stems end in the consonant й, which follows a vowel. When we add endings to these stems, the й disappears, combining with the vowel of the endings. This creates some familiar spelling changes: й + у → ю, and й + е → е.

Click the audio button to listen to the conjugated forms read aloud for you. Repeat after the speaker as you read through the chart.

Stem	Endings	Conjugated Forms	Meanings
чита́й-	+ у	→ я чита́ю (й + у)	I read, I am reading, I do read
	+ ешь	→ ты чита́ешь	You read, etc.
	+ ет	→ он/она́ чита́ет	He/she reads, etc.
	+ ем	→ мы чита́ем	We read, etc.
	+ ете	→ вы чита́ете	You read, etc.
	+ ут	→ они́ чита́ют (й + у)	They read, etc.

Most of the other verbs that you have encountered so far in this unit follow the same conjugation pattern as читай-. Their stems and basic meanings are listed below.

слу́шай = listen to	игра́й = play	де́лай = do, make
зна́й = know	отдыха́й = rest, relax	убира́й = tidy up, clean up
ду́май = think	гуля́й = stroll, fool around	рабо́тай = work

Nemnogo o jazyke: Information on stress shifts in different conjugations

Stress in Russian verb conjugation

In many Russian verbs you will learn stress will remain on the same syllable through the whole conjugation. That syllable can be in the stem (like: слу́шай-) or the ending (like жив-ˊ). However, for a small number of verbs (like пиш-) the stress in the я form is on the ending, and then the stress **shifts backwards** onto the stem in all remaining forms.

Listen to the recording of the three verbs conjugated as you read through the table and pay attention to their stress patterns.

	Always stressed on the stem	**Always stressed on the ending**	**Stress shifts backward from ending to stem**
я	слу́шаю	живу́	пишу́
ты	слу́шаешь	живёшь	пи́шешь
он/она́	слу́шает	живёт	пи́шет
мы	слу́шаем	живём	пи́шем
вы	слу́шаете	живёте	пи́шете
они́	слу́шают	живу́т	пи́шут

Appendix C: Raw Proficiency Test Scores

Table 16
Listening Comprehension Task Scores Items 1-8

	1	2	3	4	5	6	7	8
Participant								
3000	1	1	1	1	1	1	1	1
3001	1	1	1	1	1	1	1	1
3002	1	1	1	1	1	1	1	1
3003	1	1	1	1	1	1	1	1
3004	1	1	1	1	1	1	1	1
3005	1	1	1	1	1	1	1	1
3006	1	1	1	1	1	1	1	1
3007	1	1	1	1	1	1	1	1
3008	1	1	1	1	1	1	1	1
3009	1	1	1	1	1	1	1	1
3010	1	1	1	1	1	1	1	1
3011	1	1	1	1	1	1	1	1
3012	1	1	1	1	1	1	1	1
3013	0	1	1	1	1	1	1	1
3014	1	1	1	1	1	1	1	1
3015	1	1	1	1	1	1	1	1
3016	1	1	1	1	1	1	1	1

Table 17
Listening Comprehension Task Scores Items 9-16

	9	10	11	12	13	14	15	16
Participant								
3000	1	1	1	1	1	1	1	1
3001	1	1	1	1	1	1	1	1
3002	1	1	1	1	1	1	1	1
3003	1	1	1	1	1	1	1	1
3004	1	1	1	1	1	1	1	1
3005	1	1	1	1	1	1	1	1
3006	1	1	1	1	1	1	1	1
3007	1	1	1	1	1	1	1	1
3008	1	1	1	1	1	1	1	1
3009	1	1	1	1	1	1	1	1
3010	1	1	1	1	1	1	1	1
3011	1	1	1	1	1	1	1	1
3012	1	1	1	1	1	1	1	1
3013	1	1	1	1	1	1	1	1
3014	1	1	1	1	1	1	1	1
3015	1	1	1	1	1	1	1	1
3016	1	1	1	1	1	1	1	1

Table 18
Written Production Task Scores Items 1-8

	1	2	3	4	5	6	7	8
Participant								
3000	1	1	1	0	1	1	0	0
3001	1	1	1	0	1	1	1	1
3002	1	1	1	0	1	1	1	0
3003	1	1	1	0	1	1	1	1
3004	1	1	1	0	1	1	1	1
3005	1	1	1	0	1	1	1	1
3006	0	1	1	0	1	1	1	1
3007	1	1	1	0	1	1	1	1
3008	0	1	1	0	1	1	0	0
3009	0	1	0	0	1	1	0	0
3010	1	1	1	0	1	0	1	1
3011	0	1	0	0	1	1	1	0
3012	1	1	1	0	1	1	1	1
3013	0	1	1	0	1	0	1	0
3014	1	1	1	0	1	1	1	1
3015	0	1	1	0	1	1	1	0
3016	0	0	0	0	1	0	1	0

Table 19
Written Production Task Scores Items 9-16

	9	10	11	12	13	14	15	16
Participant								
3000	1	0	1	1	1	1	0	0
3001	1	1	1	1	1	1	1	0
3002	1	1	1	1	1	1	1	0
3003	1	1	1	1	1	1	0	0
3004	1	1	1	1	1	1	0	1
3005	1	1	1	1	1	1	0	1
3006	1	1	1	1	1	1	1	1
3007	1	1	1	1	1	1	1	1
3008	1	1	0	1	1	1	1	0
3009	1	0	0	1	1	1	0	0
3010	1	0	0	1	1	1	0	1
3011	1	1	0	1	1	1	1	0
3012	1	1	1	0	1	1	1	1
3013	1	1	1	0	1	1	0	0
3014	1	1	1	1	1	1	1	1
3015	1	1	1	1	1	1	1	0
3016	1	0	0	1	1	1	0	0

Table 20
Total Scores by Sub-Task

	Listening Comprehension N=16	Written Production N=16	Total N=32
Participant			
3000	16	10	26
3001	16	14	30
3002	16	13	29
3003	16	13	29
3004	16	14	30
3005	16	14	30
3006	16	14	30
3007	16	15	31
3008	16	10	26
3009	16	7	23
3010	16	11	27
3011	16	10	26
3012	16	14	30
3013	15	9	24
3014	16	15	31
3015	16	12	28
3016	16	6	22

Appendix D: Written Production Generalizations

Table 21
Written Production Generalizations Items 1-8

	1	2	3	4	5	6	7	8
Participant								
3000	aj	aj	aj	a**	aj	aj	i*	aj**
3001	aj	aj	aj	a**	aj	aj	i	i
3002	aj	aj	aj	a**	aj	aj	i	i*
3003	aj	aj	aj	i**	aj	aj	i	i***
3004	aj	aj	aj	i**	aj	aj	i	i
3005	aj	aj	aj	i**	aj	aj	i	i
3006	aj*	aj	aj	a**	aj	aj	i	i
3007	aj	aj	aj	a**	aj	aj	i	i
3008	X	aj	aj***	a**	aj	aj***	aj**	i*
3009	X	aj	X	X	aj	aj	X	i*
3010	aj	aj	aj	i**	aj	i**	i	i
3011	aj*	aj	aj*	**	aj	aj	i	i*
3012	aj	aj	aj	i**	aj	aj	i	i
3013	aj	aj	aj***	i**	aj	aj	i***	i*
3014	aj	aj	aj	a**	aj	aj	i	i
3015	i**	aj	aj	a**	aj	aj	i	aj**
3016	aj***	aj***	X	X	aj	X	aj**	X

* = Wrong person (e.g. 3rd person singular when 3rd person plural is expected)

** = Wrong verbal class
*** = Orthographic error
X = Wrong verb (e.g. “remember” instead of “walk”)

Table 22
Written Production Generalizations Items 9-16

	9	10	11	12	13	14	15	16
Participant								
3000	aj	i*	aj	aj	aj	aj	aj**	aj**
3001	aj	i	aj	aj	aj	aj	a	aj**
3002	aj	i	aj	aj	aj	aj	a	i*
3003	aj	i	aj	aj	aj	aj	i**	aj**
3004	aj	i	aj	aj	aj	aj	i**	i
3005	aj	i***	aj	aj	aj	aj	i**	i
3006	aj	i	aj	aj	aj	aj	a	i
3007	aj	i	aj	aj	aj	aj	a	i
3008	aj	i	a**	aj	aj	aj***	a	aj**
3009	aj	i*	aj*	aj	aj	aj	aj*	aj**
3010	aj	aj**	i**	aj	aj	aj	aj**	i
3011	aj	i	aj*	aj	aj	aj	a	i*
3012	aj	i	aj	X	aj	aj	a	i
3013	aj	i	aj	aj***	aj	aj	aj**	aj**
3014	aj	i***	aj	aj	aj	aj	a	i
3015	aj	i	aj	aj	aj	aj	a	aj**
3016	aj	X	aj*	aj	aj	aj	X	X

* = Wrong person (e.g. 3rd person singular when 3rd person plural is expected)
** = Wrong verbal class
*** = Orthographic error
X = Wrong verb (e.g. “remember” instead of “walk”)

Appendix E: Russian Verbal Morphology Test

Comprehension Task:

Choose the picture that best describes what you hear.
Pay attention to the verb endings.

Circle one letter only

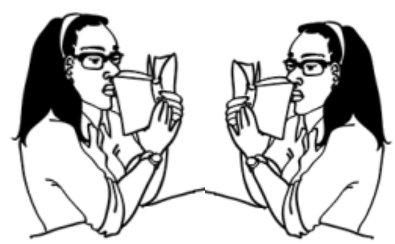
1. Профессор: Я завтракаю.
Ja zavtrakaju.
Professor: I am eating breakfast.

А. 

Б. 

2. Профессор: Я читаю.
Ja chitaju.
Professor: I am reading.

А. 

Б. 

3. Профессор: Они играют.
Oni igraju.
Professor: They are playing.

А. 

Б. 

4. Профессор: Они живут здесь.
Oni zhivut zdjes'.
Professor: They live here.

А. 

Б. 

5. Профессор: Я одеваюсь.
Ja odevajus'.
Professor: I am getting dressed.

А. 

Б. 

6. Профессор: Они плавают.
Oni plavajut.
Professor: They are swimming.

А. 

Б. 

7. Профессор: Я работаю.
Ja rabotaju.
Professor: I am working.

А. 

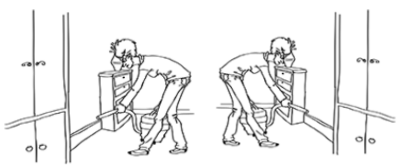
Б. 

8. Профессор: Я понимаю.
Ja ponimaju.
Professor: I understand.

А. 

Б. 

9. Профессор: Я убираю.
Ja ubiraju.
Professor: I am cleaning up.

А. 

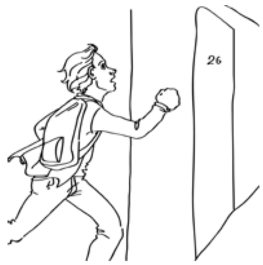
Б. 

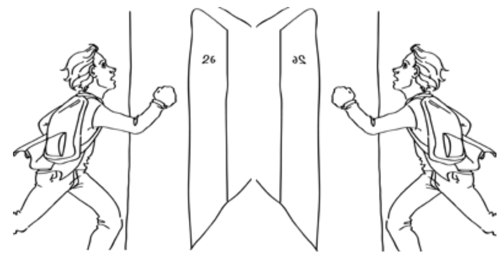
10. Профессор: Они изучают.
Oni izuchajut.
Professor: They are studying.

А. 

Б. 

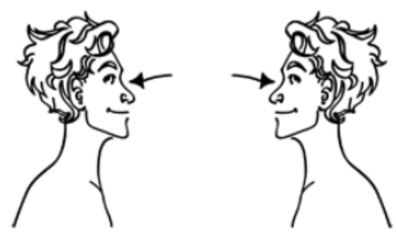
11. Профессор: Они спишут.
Oni spishut.
Professor: They are hurrying.

А. 

Б. 

12. Профессор: Я вижу.
Ja vizhu.
Professor: I see.

А. 

Б. 

13. Профессор: Они покупают.
Oni pokupajut.
Professor: They are buying.

А. 

Б. 

14. Профессор: Они пишут.
Oni pishut.
Professor: They are writing.

А. 

Б. 

15. Профессор: Они помнят.
Oni pomnjat.
Professor: They remember.

А. 

Б. 

16. Профессор: Я бегаю.
Ja bjegaju.
Professor: I am running.

А. 

Б. 

Written Production Task:

Provide the correct conjugation of the verb that best describes the picture. Pay attention to the verb endings.

1.

А: Я гуляю. Что они делают?
Ja guljaju. Chto oni delajut?
I am walking. What are they doing?

Б: Они __________.
Oni ___________.
They are ________.

Answer:

Б: Они гуляют.
Oni guljajut.

2.

А: Они отдыхают. Что ты делаешь?
Oni otdykhajut. Chto ty delajesh'?
They are resting. What are you doing?

Б: Я ________.
Ja ______.
I am ______.

Answer: Я отдыхаю.
Ja otdykhaju.
I am resting.

3.

А: Я заказываю. Что они делают?
Ja zakazyvaju. Chto oni delajut?
I am ordering. What are they doing?

Б: Они ______________________.
Oni ________________________.
They are __________________.

Answer: Они заказывают.
Oni zakazyvajut.
They are ordering.

4.

А: Я лечу домой. Что они делают?
Ja lechu domoj. Chto oni delajut?
I am flying home. What are they doing?

Б: Они ___________________.
Oni ____________________.
They are ________________.

Answer: Они летят домой.
Oni letjat domoj.
They are flying home.

5.

А: Они думают. Что ты делаешь?
Oni dumajut. Chto ty delaesh'?
They are thinking. What are you doing?

Б: Я _____________________.
Ja ___________________.
I am _________________.

Answer:
Я думаю.
Ja dumaju.
I am thinking.

6.

А: Я спрашиваю. Что они делают?
Ja sprashivaju. Chto oni delajut?
I am asking. What are they doing?

Б: Они ___________________.
Oni ____________________.
They are ___________________.

Answer: Они спрашивают.
Oni sprashivajut.
They are asking.

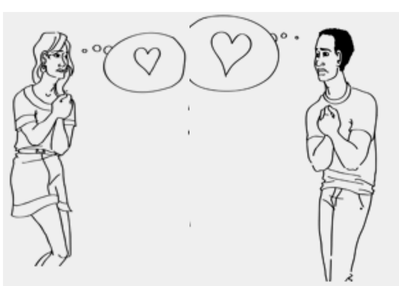
7.

А: Они говорят. Что ты делаешь?
Oni govorjat. Chto ty delajesh'?
They are talking. What are you doing?

Б: Я _____.
Ja ______.
I am ______.

Answer: Я говорю.
Ja govorju.
I am talking.

8.



А: Я его люблю. Что они делают?
Ja ego ljublju. Chto oni delajut?
I love him. What are they doing?

Б: Они его любят.
Oni ego __________.
They ________ him.

Answer:
Они его любят.
Oni ego ljubjat.
They love him.

9.

А: Они мечтают. Что ты делаешь?
Oni mechtajut. Chto ty delajesh'?
They are dreaming. What are you doing?

Б: Я _____________.
Ja _____________.
I _____________.

Answer:
Я мечтаю.
Ja mechtaju.
I am dreaming.

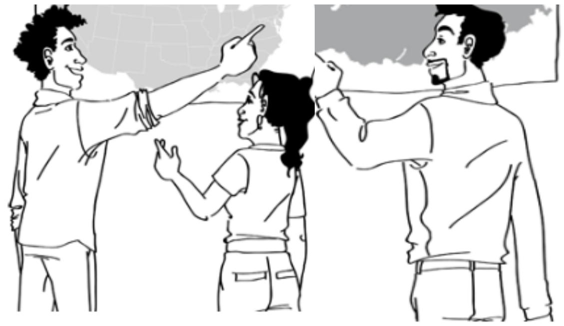
10.

А: Они дарят подарки. Что ты делаешь?
Oni darjat podarki. Chto ty delajesh'?
They are giving gifts. What are you doing?

Б: Я _______________ подарки.
Ja _______________ *podarki.*
I am _____________ gifts.

Answer:
Я дарю подарки.
Ja darju podarki.
I am giving gifts.

11.



А: Я отвечаю на вопрос. Что они делают?
Ja otvechaju na vopros. Chto oni delajut?
I am answering the question. What are they doing?

Б: Они ________________ на вопрос.
Oni ________________ na vopros.
They are ______________ the question.

Answer:
Они отвечают на вопрос.
Oni otvechajut na vopros.
They are answering.

12.

А: Они слушают музыку. Что ты делаешь?
Oni slushajut muzyku. Chto ty delajesh'?
They are listening to music. What are you doing?

Б: Я ______________ музыку.
Ja ______________ *muzyku*.
I am _____________ to music.

Answer:
Я слушаю музыку.
Ja slushaju muzyku.
I am listening to music.

13. 

А: Они делают салаты. Что ты делаешь?
Oni delajut salaty. Chto ty delajesh'?
They are making salads. What are you doing?

Б: Я _________________ салаты.
Ja _________________ salaty.
I am ______________ salads.

Answer:
Я делаю салаты.
Ja delaju salaty.
I am making salads.

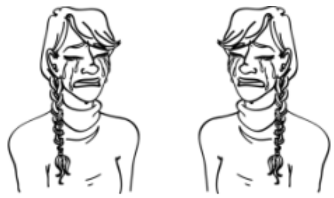
14. 

А: Они опаздывают. Что ты делаешь?
Oni opazdyvajut. Chto ty delajesh'?
They are late. What are you doing?

Б: Я __________________.
Ja _________________.
I am _______________.

Answer:
Я опаздываю.
Ja opazdyvaju.
I am late.

15.



А: Я плачу. Что они делают?
Ja plachu. Chto oni delajut?
I am crying. What are they doing?

Б: Они _____________.
Oni _______________.
They ___________.

Answer:
Они плачут.
Oni plachut.
They are crying.

16.

А: Я смотрю телевизор. Что они делают?
Ja smotrju televizor. Chto oni delajut?
I am watching. What are you doing?

Б: Они _________________ телевизор.
Oni _________________ televizor.
They are _____________ TV.

Answer:
Они смотрят телевизор.
Oni smotrjat televizor.
They are watching TV.

Table 23
Listening Comprehension Task Key

	Verb	Translation	Learned/ Novel (L or N)	Verbal Class	Inflection	Answer
1	*zavtrakat'*	to eat breakfast	L	aj	1SG	B
2	*chitat'*	to read	L	aj	1SG	A
3	*igrat'*	to play	L	aj	3PL	A
4	*zhit'*	to live	L	e	3PL	A
5	*odevat'sja*	to get dressed	N	aj	1SG	B
6	*plavat'*	to swim	N	aj	3PL	A
7	*rabotat'*	to work	L	aj	1SG	B
8	*ponimat'*	to understand	L	aj	1SG	B
9	*ubirat'*	to clean up	L	aj	1SG	B
10	*izuchat'*	to study	L	aj	3PL	B
11	*speshit'*	to hurry	N	i	3PL	B
12	*videt'*	to see	N	i	1SG	A
13	*pokupat'*	to buy	N	aj	3PL	B
14	*pisat'*	to write	L	a	3PL	A
15	*pomnit'*	to remember	N	i	3PL	A
16	*begat'*	to run	N	aj	1SG	

Table 24
Written Production Task Key

	Verb	Translation	Learned/ Novel (L or N)	Verbal Class	Inflection
1	*guljat'*	to walk	L	aj	3PL
2	*otdykhat'*	to rest	L	aj	1SG
3	*zakazyvat'*	to order	N	aj	1SG
4	*letet'*	to fly	N	i	3PL
5	*dumat'*	to think	L	aj	1SG
6	*sprashivat'*	to ask	N	aj	3PL
7	*govorit'*	to talk	L	i	1SG
8	*ljubit'*	to love	L	i	3PL
9	*mechtat'*	to dream	N	aj	1SG
10	*darit'*	to give	N	i	1SG
11	*otvechat'*	to answer	N	aj	3PL
12	*slushat'*	to listen	L	aj	1SG
13	*delat'*	to do/make	L	aj	1SG
14	*opazdyvat'*	to be late	N	aj	1SG
15	*plakat'*	to cry	N	a	3PL
16	*smotret'*	to watch	L	i	3PL