REPRESENTATIONAL AND PROCESSING CONSTRAINTS ON THE ACQUISITION OF CASE AND GENDER BY L1 ENGLISH LEARNERS OF RUSSIAN: A CORPUS STUDY

by

Gina M. Peirce

BA, The College of Wooster, 1995

MA, University of Michigan, 1997

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Master of Arts in Linguistics

University of Pittsburgh

2015

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Gina M. Peirce

It was defended on

August 3, 2015

and approved by

David J. Birnbaum, PhD, Professor and Chair, Department of Slavic Languages and Literatures

Na-Rae Han, PhD, Lecturer, Department of Linguistics

Thesis Advisor: Alan Juffs, PhD, Professor, Department of Linguistics

REPRESENTATIONAL AND PROCESSING CONSTRAINTS ON THE ACQUISITION OF CASE AND GENDER BY L1 ENGLISH LEARNERS OF RUSSIAN: A CORPUS STUDY

Gina M. Peirce, MA

University of Pittsburgh, 2015

Inflectional morphology errors have been found to be prevalent in adult second language production, even by learners with high overall proficiency and extensive L2 exposure. Particular difficulties are observed with acquisition of complex inflectional systems by learners whose first languages are less morphologically rich. Thus, L1 English learners acquiring the Russian case and gender systems can provide a testing ground for competing approaches to L2 morphological acquisition. This study utilizes the Russian Learner Corpus of Academic Writing, consisting of texts by advanced learners in Portland State University's Russian Language Flagship program, to compare frequencies of case and gender-marking errors in timed versus untimed compositions by a sample of heritage and traditional L2 learners in order to gauge the putative effect of processing pressure on such errors.

It was predicted that significantly more frequent errors in timed compositions would support the position that inflectional morphology errors by advanced learners largely reflect processing difficulties under time pressure. However, results for the heritage group showed that while descriptive tendencies appeared to point toward processing difficulties, the difference between timed and untimed error rates did not reach statistical significance. In the L2 group,

error rates were higher in students' *untimed* texts than in their timed texts, although this difference also did not reach significance.

The lack of reliable differences between timed and untimed error rates could be interpreted as demonstrating representational deficits in learners' interlanguage grammar, particularly in the L2 group. However, the greater complexity (measured by words per T-unit) of the untimed essays provides an alternative explanation for the unexpected finding of a higher untimed error rate among the L2 learners, for whom the correlation between differences in complexity and error rates for timed and untimed texts approached significance.

For the heritage learners, error rates appeared to be affected more by time pressure than text complexity. In addition, the heritage group had lower case-marking and significantly lower gender-marking error rates than the L2 group. This finding suggests that heritage learners are less likely than traditional L2 learners to show evidence of possible representational deficits of nominal functional features in their interlanguage grammar.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Special thanks are extended to Olesya Kisselev (Department of Applied Linguistics, Pennsylvania State University) for providing the author of this thesis with access to and explanation of the RULEC dataset, and to Yerzhan Shyngayev (Department of Linguistics, University of Pittsburgh) for annotating the case and gender-marking errors with XML tags in the learner corpus texts used for this study.

I. INTRODUCTION

Second language acquisition (SLA) researchers have proposed various explanations for the prevalence of inflectional morphology errors in the production of adult L2 learners, including learners with a high overall proficiency level and many years of exposure to the L2. These involve such diverse causes as L1 influence (e.g., Portin, et al., 2007 and 2008); online processing difficulties resulting from high cognitive load (e.g., Prévost and White, 2000; McDonald, 2006); and the weakness of morphosyntactic cues to form-meaning mapping (e.g., Kempe and MacWhinney, 1998; Taraban and Kempe, 1999). Some scholars posit irreparable representational deficits in L2 learner competence due either to mismatched grammatical categories in the L1 and L2, with L2 speakers lacking full access to Universal Grammar (e.g., Hawkins and Chan, 1997), or to adult learners' lack of access to language-specific implicit learning mechanisms (e.g., Clahsen, et al., 2010). Particular difficulties are observed with acquisition of complex systems of inflectional morphology, as in Russian and other Slavic languages, by learners whose native languages are less morphologically rich. Thus, the experiences of L1 English learners acquiring the Russian case and gender systems can provide a testing ground for competing approaches to L2 morphological processing.

This thesis will begin by reviewing studies of acquisition of inflectional morphology, focusing on the following questions: (1) Are errors in case and gender marking by L1 English learners of Russian mainly performance errors caused by processing difficulties and/or weak morphosyntactic cues, or competence errors caused by representational deficits? (2) Does the

absence of morphological case and gender marking on nouns and adjectives in English prevent such learners from ever fully acquiring these forms in Russian, or can they eventually be acquired through means such as exposure to sufficient input, explicit instruction and/or conscious self-monitoring?

The thesis will then present an analysis of the Russian Learner Corpus of Academic Writing (RULEC), a collection of texts written by advanced L2 and heritage learners in the Russian Language Flagship program at Portland State University. This analysis will compare the frequency of case and gender-marking errors in timed versus untimed compositions by the same groups of L2 and heritage learners in the RULEC corpus. Selection of this research design was based on the premise that significantly more frequent errors in the timed than the untimed compositions would suggest that instances of non-target-like inflectional morphology produced by advanced language learners are largely performance errors, reflecting production processing difficulties under time pressure. Conversely, a lack of significant differences in error rates would support the view that persistent difficulties with inflectional morphology could be primarily competence errors, which may be caused by representational deficits. Interpretation of the actual results of the corpus data analysis, however, presents a more complex set of issues; these will be detailed in the discussion section of this paper.

II. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

The research questions addressed here are situated in a framework of competing theoretical models of the target of acquisition itself (theories of morphology), as well as theories of processing and acquisition of inflectional morphology. Existing models may be grouped into two main categories, as outlined in Gor (2010a): dual-system approaches, which make a categorical distinction between rule-based learning of regular inflections and associational learning of irregular inflections; and single-system approaches, which are connectionist models based entirely on associational learning from input, resulting in storage of whole words in memory rather than decomposition into separate morphemes for processing. Under the single-system model, there is no morphological level of representation and processing that is distinct from phonological and semantic processing, as all inflected words are processed by an associative patterning mechanism. The dual-system model, however, posits a separate mechanism of symbolic rule computation by which regular inflected forms are processed.

Gor noted that languages such as Russian are problematic for the dual-system model: "In languages with rich inflectional morphology, there is no sharp division between regular and irregular inflection but, rather, several inflectional patterns ranging in regularity" (Gor, 2010a, p. 5). Thus, Gor concluded that a categorical distinction cannot be universal and proposed a continuum between regular and irregular processing, suggesting cross-linguistic differences exist among languages with varying morphological richness. This "rules and probabilities" model, proposed earlier by Jackendoff (2002), was supported by Gor and Cook's study of production

and recognition of conjugated verbs by Russian L2 and heritage learners (2010b). Both groups' choices of inflectional patterns were found to depend on "implicit knowledge of probabilities and efficiency in the retrieval of morphological cues to the inflectional pattern, and application of complex morphophonological rules" (p. 118), although heritage speakers relied on whole-word storage to a greater degree than the L2 learners. Similar findings were reported in Gor and Vdovina (2010c). A masked priming study of L1 Russian speakers by Kazanina, et al. (2008) also departed from both dual-system and single-system models, suggesting that decomposition of multimorphemic words occurs automatically at an early stage of lexical processing and "in languages with rich morphological structure, such as Russian, an early automatic decomposition may have a stronger weight than in languages with limited use of morphology" (p. 814).

A usage-based approach proposed by scholars such as Dabrowska (2008) combines elements of the dual-system and single-system models, explaining morphological processing through schemas that abstract rules from patterns in input. Tkachenko and Chernigovskaya (2010) also supported a usage-based account, finding that L2 acquisition of Russian verbal inflection is affected by both type and token frequency. In addition, a dual-route model claims that morphological decomposition and whole-word access are simultaneously activated to maximize processing speed (Gor, 2010a, p. 6). Alegre and Gordon (1999) argued that frequency effects in processing do not exclude the possibility of a rule-based system, based on lexical decision experiments showing "there are indeed whole-word representations available for regularly inflected items," but "this is not true for items in the lower end of the frequency distribution," which must be processed through morphological decomposition (p. 57).

Competing models of the basic mechanisms of morphological processing are accompanied by diverse theories on differences between L1 and L2 acquisition. For instance,

both Clahsen, et al. (2010) and Jiang (2004) have posited representational deficits in inflectional grammar of adult L2 learners. Clahsen, Felser and colleagues extended their Shallow Structure Hypothesis from sentence processing to morphosyntactic processing, claiming "L2 learners are less sensitive to morphological structure than L1 speakers and rely more on [lexical] storage than morphological decomposition" (Gor, 2010a, p. 10). This position is based on Clahsen's support for the dual model of processing, in which both L1 and L2 speakers exhibit qualitative differences between regular and irregular inflected forms in their grammatical systems (e.g., Clahsen, 1995).

In a review of studies of online morphological processing by advanced L2 learners, Clahsen, et al. (2010) reported that non-native speakers rely strongly on the declarative memory system rather than decomposition for processing morphologically complex words, which is consistent with M. Ullman's theory that the procedural or implicit knowledge system is attenuated in adult language learners (Clahsen, 2010, p. 38). However, this contradicts findings by Portin, et al. (2007) that initially "late L2 learners only store full-form representations of very high-frequency inflected words and use decomposition to process all other inflected words" (Gor, 2010a, p. 11), while with increasing proficiency they also develop whole-word representations of lower-frequency items. Portin proposed that with continued exposure to L2 input, learners' mental representations become more native-like over time.

Jiang (2004) showed that L1 Chinese learners of English were insensitive to morphological plural markers and number agreement in English, as demonstrated by their lack of significant differences in reading times for sentences involving number agreement and disagreement. The subjects' performance on a written test demonstrated they had explicit knowledge of the rules for number agreement in English. Thus, Jiang concluded that their

insensitivity to plural morphology in the timed, online task resulted from a deficiency in implicit mental representations, which could possibly be attributed to age-related critical period effects and the absence of plural morphology in their Chinese L1 (p. 627). In this view, the lack of a grammatical number category in these subjects' L1 required them to store English number agreement only within their explicit or declarative knowledge.

An opposing viewpoint was provided by Prévost and White's (2000) study testing the Missing Surface Inflection Hypothesis (MSIH) against the Impaired Representation Hypothesis (IRH). The MSIH posits that learners have unconscious knowledge of L2 morphological features, but experience difficulty with realization of surface forms in online processing, while the IRH suggests permanent impairment of interlanguage grammar at the level of functional categories or features in learners' mental representations (Prévost and White, 2000, pp. 108-110). A related account to the IRH is the Failed Functional Features Hypothesis or FFFH, proposed by Hawkins and Chan (1997), which holds that Universal Grammar is only partially available to adult L2 learners; thus, such learners cannot access features associated with functional categories in their L2 that have different parameter settings from their L1. Prévost and White, however, found that missing inflection in L2 French and German learners' oral production of finite verbs reflected use of non-finite verbs as default forms under online processing pressure. Because the subjects rarely made the reverse error of substituting finite for non-finite verbs, and typically inflected finite verbs correctly when these were used, their production exhibited systematicity rather than randomness. These findings supported the MSIH, as the IRH would predict arbitrary use of inflectional morphology, with finite and non-finite forms being used randomly.

Spinner and Juffs (2008) incorporated Prévost and White's findings into a longitudinal study of gender agreement errors by two L2 German learners, stating that the MSIH

characterizes such errors as “failure to select the appropriate morphological form in spellout or ‘mapping’” (p. 321). Spinner and Juffs’ subjects made fewer errors inflecting nouns with natural gender and highly frequent nouns than others, suggesting that some errors were caused by failure to associate the correct gender with specific nouns. The authors noted these lexical problems were likely related to “weakness of phonological and semantic cues [to gender] in German” (p. 344). This could also be relevant to L2 Russian learners, particularly regarding semantic cues, which most Russian nouns without natural gender lack.

Spinner and Juffs found no significant advantage for the L1 Italian over the L1 Turkish learner on acquisition of German gender marking, or for the L1 Turkish speaker on German case marking. These findings suggested the subjects’ errors were not due to the absence of relevant functional features in their native languages, contradicting the FFFH. Spinner and Juffs also found evidence of unidirectional errors, which, as in Prévost and White, indicated systematicity rather than randomness in subjects’ interlanguage and were consistent with the MSIH. The subjects’ frequent omission or defaulting of inflected forms under conditions of high cognitive load (such as long or complex phrases) also pointed toward processing difficulties. Spinner and Juffs concluded that “errors may arise from multiple sources, including lexical learning, processing, mapping, and lexical assignment/parsing the input” (p. 344). Hopp (2013) confirmed that lexical learning of gender assignment may cause errors by advanced L2 learners, suggesting “variability in the use of gender agreement in real-time comprehension reflects lower levels of activation of and access to gender nodes in the adult bilingual mental lexicon rather than representational deficits in L2 grammars” (p. 53).

Another explanation is suggested by the Feature Reassembly approach (e.g., Spinner, 2013), in which learners must detect how grammatical features are spelled out in their L2 and

reassemble the feature matrices from their L1 into the target language configuration. This is particularly difficult when features that are present in and absent from the L1 are fused into a single morphological form in the L2, such as gender and number marking in Swahili. Spinner (2013) claimed that Feature Reassembly provided a more convincing explanation than representational deficit accounts for the tendency among L1 English speakers to produce more errors in marking number than gender on Swahili nouns, even though number but not gender is overtly marked on English nouns (p. 476).

McDonald (2006) also argued against representational deficits, but from yet another perspective. She focused on cognitive load to demonstrate that processing difficulties are a likelier cause of L2 learner errors. Cognitive load theory holds that "working memory limitations must be taken into account" to analyze learning of "biologically secondary knowledge," which is not essential to normal human functioning and takes conscious effort to acquire (Sweller, et al., 2011, pp. 52-53), such as adult SLA. Recently learned information that has not been transferred to long-term memory must be stored in working memory; yet working memory capacity is severely limited. Through grammaticality judgment tests with L1 English speakers and adult L2 learners of English, McDonald found that L2 learners in a baseline unstressed condition performed similarly to native speakers in a noise stress condition or under increased cognitive load. She concluded that "while late L2 learners perform quite differently from unstressed natives, this is not adequate evidence to say their grammatical knowledge is qualitatively different. Rather, it is possible […] that late L2 learners actually have a large overlap in grammatical knowledge with native speakers; they are just processing the sentences under difficult conditions" (p. 397).

Two studies by Portin, et al. provided further evidence that adult L2 learners could develop mental representations similar to those of native speakers, while highlighting the role of L1 effects on acquisition of L2 morphology. Portin (2007) found that L1 Finnish speakers primarily used morphological decomposition in visual lexical decision tasks involving low-frequency inflected Swedish nouns, while accessing full-form representations of medium and high-frequency Swedish nouns, as demonstrated by greater differences between response times for inflected and monomorphemic test items on low-frequency words. The authors concluded that "lengthy formal study […] can provide a late learner with nativelike full-form input representations" (p. 151). This finding contradicts Clahsen's (2010) position that L2 learners rely more on full-form representations of inflected words than native speakers. Portin (2008) explored whether L1 Hungarian and Chinese speakers would transfer strategies from their native languages for processing inflected Swedish nouns, and the results suggested that such L1 influence did occur. The Hungarian group had longer response times for inflected than monomorphemic test items on low and medium-frequency Swedish words, while the Chinese group had similar response times for inflected and monomorphemic test items regardless of frequency, indicating they processed all inflected words through full-form representations. Portin attributed this to the fact that Swedish is less highly inflected than Hungarian but has more complex morphology than Chinese, in which inflection is virtually absent.

Thus, straightforward accounts of L2 learners moving toward greater use of either decomposition, as in Clahsen (2010), or full-form representations, as in Portin (2007), with increasing proficiency appear to be incomplete. Despite claims that L1 effects are absent from L2 processing, e.g. by Silva and Clahsen (cf. Rehak and Juffs, 2011, p. 126), Portin (2008) presents compelling evidence that the structures of a learner's particular L1 and L2 must be

considered. As indicated in Gor (2010a) and Kazanina (2008), the relative importance of decomposition and full-form storage may vary with morphological richness of specific languages. Processing of full-form representations is more viable for languages with fewer inflected forms, in which this strategy places less strain on learners' limited memory capacity. A morphologically complex language, on the other hand—such as Russian, with its three grammatical genders, six cases, and extensive verbal conjugation paradigms—would require storage of a large number of inflected forms of each noun, adjective and verb, thus placing greater pressure on cognitive resources.

Others have argued for the possibility of native-like representations among L2 learners through a connectionist approach. Kempe and MacWhinney (1998) worked from the Competition Model, which "focuses on inflections as cues to underlying thematic roles and pragmatic functions" (p. 545), to hypothesize that stronger case marking cues lead L1 English learners of Russian to acquire case marking more rapidly than learners of German. Data from a picture-choice task confirmed that Russian learners made fewer errors on case-marked sentences with OVS word order than German learners. Kempe and MacWhinney concluded their evidence for input-based associative learning indicates "it is reasonable to expect that the comprehension strategies of advanced L2 learners should resemble those of native speakers" (p. 568). This view conflicts with representational deficit accounts such as Clahsen (2010) and Jiang (2004), but is supported by other studies such as Sagarra and Herschensohn (2010), which concluded from online and offline tests of morphological processing by L1 and L2 Spanish speakers that "adult learners with a certain proficiency level can demonstrate grammatical knowledge and implementation that is qualitatively comparable to that of native speakers" (p. 2035).

Taraban and Kempe (1999) presented another connectionist account of gender processing, predicting that L1 and advanced L2 Russian speakers would utilize a common learning mechanism to infer the gender of unfamiliar nouns. They found that ambiguous phonological cues to gender marking led to slower and less accurate sentence processing by both L1 and L2 subjects on a self-paced reading task, while the presence of adjectives reliably marked for gender improved performance for both groups on sentences with phonologically ambiguous nouns. The authors concluded, "L1 and L2 speakers may rely on similar learning mechanisms for mastering gender and simply may be at different points on the learning curve" (p. 144).

Other research on morphological processing has focused on heritage speakers, a group that does not fit neatly into an L1/L2 dichotomy. Polinsky's (2008) experiments with Russian heritage speakers showed systematicity in their processing of gender marking, despite differences from monolingual Russian speakers. When the subjects were asked to supply forms of adjectives and possessive pronouns that agreed grammatically with Russian nouns and to judge the acceptability of adjective-noun combinations, their accuracy approached that of a monolingual control group for masculine nouns only. Polinsky concluded "the category of gender is still fully represented" in the interlanguage grammar of Russian heritage speakers, "but its actual structuring undergoes significant reanalysis. While gender assignment principles manifested in agreement are different from the baseline, they are definitely not random" (p. 55). Thus, Polinsky did not view heritage speakers as experiencing representational deficits, although their acquisition of Russian grammar was incomplete due to insufficient input and instruction. She recommended further research comparing heritage speakers with L2 learners, observing that "such a comparison would allow us to identify those properties of noun categorization and gender priming that cut across all these groups and those that are specific to L1, L2, and to

heritage speakers only" (p. 64). This thesis addresses the need for such data by comparing evidence for representational and processing difficulties with inflectional morphology among the Russian heritage and L2 learner groups in the RULEC corpus.

In conclusion, while some L2 morphological processing studies have found evidence for representational deficits through online tasks such as self-paced reading (e.g., Jiang, 2004) or masked priming (e.g., Silva and Clahsen as summarized in Rehak and Juffs, 2011), several others have presented persuasive evidence that L2 errors in inflectional morphology may be due to other causes, such as increased processing load in timed versus untimed tasks and lexical learning of gender assignment. Such findings have resulted from studies utilizing both oral production data (e.g., Prévost and White, 2000; Spinner and Juffs, 2008; and Polinsky, 2008), and input processing tasks such as lexical decision (e.g., Portin, et al., 2007 and 2008), grammaticality judgment (e.g., McDonald, 2006), and self-paced reading (e.g., Taraban and Kempe, 1999). This array of evidence suggests that representational deficit approaches have not adequately accounted for L2 learner difficulties with inflectional morphology.

Additionally, this review indicates that the relatively few existing studies on processing of Russian case and gender by L1 English adult learners (e.g., Kempe and MacWhinney, 1998; Taraban and Kempe, 1999) have examined this phenomenon mainly from a connectionist perspective. A need is present for studies testing other models of acquisition of Russian case and gender by L1 English learners to explain the prevalence of production and comprehension errors. Approaches based on processing difficulties associated with working memory limitations and cognitive load, as well as those focusing on transfer of morphological processing strategies from the L1, have shown promise in existing literature and could benefit from testing with L1 English learners of the complex Russian system of nominal morphology.

This thesis will contribute to addressing these needs, along with that identified by Polinsky (2008) for comparison of heritage speakers and traditional L2 learners. Toward these ends, the corpus study detailed below will investigate whether a sample of advanced Russian language learners made significantly more case and gender-marking errors in compositions written under time pressure, which could lead to increased production processing difficulties, than in untimed writing assignments. This analysis will also explore whether different patterns can be observed in the effects of time pressure on morphological accuracy of written production by Russian heritage versus non-heritage learners.

III. METHODOLOGY

All data for this study were obtained from RULEC, the Russian Learner Corpus of Academic Writing. The use of learner corpora is a somewhat recent development in SLA research, but such corpora can provide a rich source of data for discerning patterns in the interlanguage grammar of L2 learners.[1] The Russian Language Flagship program faculty at Portland State University who designed and assembled RULEC sought specifically to address "the relative lack of empirical research on advanced interlanguage" (Alsufieva, et al., 2012, p. 85). Both texts written during classes with a time limit and those written at home without a time limit were deliberately included in the corpus to facilitate comparisons of language learner production under timed and untimed conditions, as the RULEC designers recognized time limits as "one of the most important variables influencing the accuracy and complexity of writing" (p. 87). The opportunity afforded by a learner corpus to directly compare the same type of linguistic output (written production) under conditions of processing pressure imposed by a time limit, versus the lack of such pressure in an offline task without a time limit, makes a corpus study an appropriate methodology for investigating the questions about L2 difficulties with inflectional morphology raised in this thesis. It is also a useful methodology for comparison of heritage speakers and traditional L2 learners, as anticipated by the RULEC designers, who viewed the corpus as "a tool

[1] While corpus data can show production performance and thus enable some inferences regarding learner competence, it is recognized that such data do not reveal how interlanguage grammars may be constrained (that is, what learners know not to be possible in their L2). For this reason, production data should ideally be used together with other sources in SLA research; but this is beyond the scope of the current exploratory study.

that may help uncover universal or group-specific patterns of Russian language acquisition and build profiles of various groups of RFL [Russian as a Foreign Language] learners" (p. 89).

The RULEC corpus consists of approximately 3,800 Russian learner texts composed over four years, ranging in length from single paragraphs to full research papers and containing a total of about 750,000 words. The texts were written as course assignments by 36 students enrolled in Portland State University's Russian Language Flagship program, including 17 foreign language learners whose L1 is English, as well as 19 heritage learners who were either born in the U.S. to Russian-speaking parents or moved to the U.S. with their families during childhood ("RULEC: Russian Learner Corpus of Academic Writing," 2009). The federally funded Russian Language Flagship offers a multi-year curriculum of intensive instruction, content-based language courses in students' major fields, and extended study abroad experiences to develop professional-level proficiency ("Language Flagship: Creating Global Professionals," 2013).

All students had been tested at a Russian oral proficiency level between Intermediate Mid and Advanced High on the ACTFL scale at the time when they wrote the texts. While ACTFL Oral Proficiency Interviews are not designed to score low-level morphological errors, this testing serves to establish similar overall proficiency levels between the two groups of students under examination, thus facilitating comparison of data from the L2 and heritage learner groups in this study. This comparison was employed to test the possibility that heritage students' exposure to Russian language input in early childhood results in different patterns of morphological processing than are found among traditional L2 learners, as suggested by Polinsky (2008) and Gor (2010b).

Each text in the corpus is tagged according to criteria including the L2 or heritage status and overall proficiency level of the author, timed or untimed writing assignment, text type and

function, etc. For this study, a subset of texts was selected for analysis that was evenly divided between those written by L2 and heritage learners, and subdivided within each of these two categories into timed and untimed compositions. The texts chosen in each of the four combinations of categories (L2/timed, heritage/timed, L2/untimed, and heritage/untimed) were matched to the greatest possible extent on all other criteria, including text length, to eliminate confounding variables. A total of 240 texts by 24 authors (12 L2 and 12 heritage learners, with each author contributing 5 timed and 5 untimed texts) were selected for the study. The number of texts used for analysis was subsequently reduced to 232, as some duplicates were found in the originally selected sample. These 232 texts contain a total of 25,741 words and 2,226 T-units (sentences or independent clauses with any accompanying dependent clauses). The average text length is approximately 111 words.

While this study includes fewer subjects than the 15 or more per group typically recommended for comparative studies in applied linguistics (Dörnyei, 2007, pp. 99-100), the use of a repeated measures design with multiple compositions by the same subjects was expected to accommodate a smaller sample size by reducing the role of variation in learner error rates resulting from individual differences among subjects. In addition, the version of RULEC currently available to researchers includes texts from only 28 (13 heritage and 15 L2 learners) of the 36 students whose compositions were originally collected, because the corpus designers viewed discrepancies among the quantities of texts from individual students as excessive, and thus reduced the number of subjects to provide more nearly equal representation of the remaining authors in RULEC (O. Kisselev, personal communication, October 21, 2014).

A native Russian-speaking graduate student in the University of Pittsburgh's Department of Linguistics identified and tagged all inflectional morphology errors on nouns, adjectives and

determiners in the texts to be analyzed. Using oXygen XML Editor software, tags were applied to each error to show the correct grammatical case, gender and number and to indicate whether the author used an incorrect case, gender and/or number, or whether the type of error could not be determined.[2] The annotation system used for error tagging is presented in Appendix A, and an example of a text annotated for this study is shown in Appendix B.

Because Russian is a synthetic language with a single inflectional ending to mark case, gender and number, and also due to frequent syncretism of inflectional endings—for instance, the same feminine endings are used in several oblique cases, and masculine and neuter endings are identical to one another in several cases—it is not always possible to distinguish case errors from gender or number errors or to identify the exact case, gender and number that the author attempted to use on a particular word. For example, in one text, the author wrote "в генетики" * ["v genetiki" / "in genetics"] in place of the correct feminine singular prepositional case form, "в генетике" ["v genetike"]. This could be a case error if the student was attempting to use genitive instead of prepositional case, or both a number and case error if the student was attempting to use a plural accusative case form. Of the 493 total annotated errors in the sample of texts used for this study, 27 errors or 5.5% were tagged as being of an unclear type.

To verify the accuracy of error annotation, the author of this thesis checked the tagging of 20% of the selected texts, including one timed and one untimed text by each of the 24 student authors. Among the 106 words annotated with error tags in these 48 texts, differences in rater judgments concerning whether or not an item constituted a nominal morphology error were present for 6 tags, while differences in categorization of the type of error (e.g., case error, gender

[2] Tagging to indicate correct cases will allow for analysis in future studies of factors other than the timed or untimed status of compositions that may affect error frequencies, such as the possibility that learners make fewer errors on marking structural cases (nominative and accusative) than inherent or lexical cases. However, analysis of these factors is beyond the scope of this thesis.

error, or morphological error of unclear type) were present for another 9 tags. This constituted a 94.3% inter-annotator agreement rate on the overall number of nominal morphology errors, and an 85.8% agreement rate on both the presence and categorization of such errors.

After the selected texts had been tagged, the frequency of errors in nominal morphology was calculated for each category of texts. The independent variables in this study were the timed or untimed status of the texts (a within-subjects variable) and the L2 or heritage language background of the authors (a between-subjects variable), while the dependent variable was the frequency of nominal morphology errors in the texts. Error frequency was measured as the rate of errors per T-unit; the number of T-units in each text was manually counted and recorded by the annotator during the error tagging process. In addition to overall error counts, this study used a Python script (see Appendix C) to calculate error totals for each category: case, gender, number, and morphological errors of indeterminate type, which—due to syncretism as mentioned above—could potentially be errors of case, gender and/or number. Errors judged during the tagging process to be orthographic or phonological, such as using an allomorph associated with "hard" adjectives where an allomorph associated with "soft" adjectives (those with stems ending in palatalized [nʲ]) is required, were not annotated as inflectional morphology errors and thus are excluded from all error frequency calculations for this study.

Statistical analysis of the corpus data consisted of repeated measures ANOVA testing to compare the rates of morphological errors within each group of subjects (L2 and heritage learners) in the timed versus untimed conditions, as well as between the subject groups in each condition, thus investigating variance among the error rates in all four combinations of group and condition. In addition, based on the results of the repeated measures analysis, correlations were

tested between the complexity of the learner texts (measured as the number of words per T-unit) and error rates in the timed and untimed compositions of each subject group.

IV. PREDICTIONS

The methodology for this study utilized a corpus of written production data by L2 learners and thus differed from the morphological processing studies summarized in the literature review, which used methodologies including self-paced reading, masked priming, grammaticality judgment, and other input processing measures. However, based on the evidence presented in those studies, it was hypothesized prior to analysis of the RULEC data that the following findings would emerge from this study.

First, on the question of representational deficits in L2 grammatical categories that are not present in the L1, available evidence suggests that the absence of overt case and gender marking on nouns and adjectives in a learner's L1 (as in English, which has other inflectional morphemes to mark number, possession, etc.) does not prevent acquisition of these forms in a more morphologically complex L2, given sufficient input. On the other hand, the near absence of any inflectional morphology in a learner's L1 (as in Chinese) does appear to raise barriers to native-like processing, or at least to the availability of morphological decomposition as a processing strategy. Such barriers are not necessarily due to irreparable representational deficits, however. Another possibility is that large differences between the morphological structures of the L1 and L2, which may be combined with low levels of working memory in some individuals, can lead to increased processing difficulties that are evidenced by slower processing and more errors in the production and comprehension of L2 inflectional morphology.

Thus, it was predicted that the RULEC corpus study would reveal significantly less frequent case and gender errors by Russian language learners in their untimed written compositions (an offline task) than in their timed compositions (which may be viewed as an online task, in the sense that the subjects performed it under time pressure). This finding would support the Missing Surface Inflection Hypothesis (MSIH), which predicts that all L2 features can be represented in learners' interlanguage grammar,[3] even if they may not be fully realized in online production and comprehension due to processing difficulties. Alternatively, if no significant differences were found between the timed and untimed compositions in the frequency of case and gender errors, this result would suggest that the production of non-target-like inflectional morphology by advanced L2 learners may reflect competence errors caused by representational deficits. The latter finding could support either Clahsen's model, which posits impaired mental representations in the grammar of all L2 learners with no L1 effects, or the Failed Functional Features Hypothesis (FFFH), which predicts deficits in L2 learners' representations of grammatical categories that are not present in their L1.

On the question of dual-system and single-system morphological processing models, the evidence presented in the literature review suggests that universally there may not be two separate processing systems based on a categorical distinction between regular and irregular morphology (since this distinction cannot be clearly made in Slavic languages, for example), but there are two possible processing routes based on morphological decomposition and full-form representations of inflected words. Factors influencing the route used for a given processing task may include morphological complexity of the L1 and L2, with full-form storage playing a larger role in languages with fewer inflected forms; type frequency and neighborhood density of

[3] It is recognized that the argument that untimed monitored output may reflect only "learning" or conscious knowledge, rather than "acquisition" or underlying knowledge of the L2 (e.g., Krashen, 2009), poses a potential problem for this study.

particular lexical items in the L2; learner proficiency levels in the L2; and individual differences in working memory. This hypothesis could not be directly tested in a corpus study, as output processing in a written task is not a suitable methodology to measure learner processing time for individual words, which requires the use of an input processing task such as masked priming, lexical decision, etc.

It was also anticipated that this study would reveal greater differences in case and gender-marking error rates between timed and untimed compositions by the Russian heritage speakers than the L2 learners whose writing samples are included in RULEC. Due to their extensive exposure to morphologically rich Russian language input in early childhood, as well as the fact that Russian is at least chronologically their first language, it was hypothesized that heritage speakers may be less likely than L2 learners to exhibit evidence of representational deficits and more able to use decomposition as a strategy for acquiring complex inflectional morphology. Thus, a greater proportion of the case and gender errors made by heritage speakers than by L2 learners of Russian may be attributable to processing difficulties, which would be reflected in larger differences in error rates between timed and untimed texts for the heritage learners.

The above predictions and statistical comparisons that are employed in this study to test these hypotheses are summarized in Table 1.

Table 1: Summary of Predictions and Tests

Research Question	**Hypothesis**	**Statistical Test**
1. Do advanced Russian language learners make more nominal morphology errors in written production under time pressure than in the absence of time pressure?	The rates of case and gender-marking errors are hypothesized to be higher in learner compositions written under time pressure than without time pressure.	Repeated measures ANOVA testing should show significant differences in nominal morphology error rates between the timed and untimed conditions.
2. Is morphological accuracy of written production affected differently by time pressure among heritage versus non-heritage learners of Russian?	It is hypothesized that there will be larger differences in case and gender-marking error rates between timed and untimed compositions for the Russian heritage speaker group than the traditional L2 learner group.	Repeated measures ANOVA testing should show a significant effect of the interaction between learner group (heritage or L2) and condition (timed or untimed) on nominal morphology error rates.

V. RESULTS

The raw numbers of nominal morphology errors and T-units found in the selected timed and untimed texts of the heritage and traditional L2 learner groups are provided in Table 2, while the overall error rates per T-unit by group and condition are provided in Table 3.

Table 2: Total Errors and T-units by Text Category

Text Category	**Total Errors**	**Total T-Units**
Heritage Group – Timed	105	511
Heritage Group – Untimed	74	516
L2 Group – Timed	146	617
L2 Group – Untimed	168	582

Table 3: Overall Error Rates per T-unit by Learner Group and Condition

Group/Condition	**Timed Condition**	**Untimed Condition**
Heritage Learner Group	0.195 (0.119)	0.134 (0.104)
L2 Learner Group	0.249 (0.160)	0.281 (0.172)

Analysis of the sample of RULEC texts annotated for this study revealed that the heritage learner group had an overall rate of 0.134 nominal morphology errors per T-unit in their untimed compositions, while their overall error rate in timed compositions was 45.5% higher at 0.195 errors per T-unit. Although this difference appears suggestive of possible processing difficulties leading to a higher error rate in the timed condition, it did not reach statistical significance.

Among the L2 learner group, the overall error rate was actually 12.9% higher in their untimed compositions (0.281 errors per T-unit) than in their timed compositions (0.249 errors per T-unit). This difference also did not reach statistical significance, but the unexpected finding of an apparently higher rate of errors in the absence of time pressure among this group of learners raises the question of why such a result might have occurred. A higher untimed error rate would be particularly surprising, given that this type of writing assignment affords students the option to use dictionaries and other resources that are unlikely to be available for timed in-class assignments. To investigate the possibility that the L2 learners may have written more complex texts in the untimed condition, thus resulting in a higher error rate, a correlation analysis was conducted to check for a potential relationship between text complexity (measured as the number of words per T-unit in the student texts) and error rates in the timed and untimed essays by each group; results of that analysis will be reported below.

The repeated measures ANOVA analysis did not reveal either a significant main effect of the within-subjects variable of timed versus untimed condition ($F(1, 22) = .227$, $p = .638$), or a significant effect of the interaction of condition and the between-subjects variable of learner group ($F(1, 22) = 2.260$, $p = .147$), on overall error rates. However, the main effect of group very closely approached significance ($F(1, 22) = 4.259$, $p = .051$), with a higher error rate for the L2 learner group (0.265 per T-unit) than for the heritage learner group (0.164 per T-unit). The observed power statistic of .505 for the main effect of group, which does not meet the generally accepted standard of .80, suggests that the relatively small sample size may have limited the statistical reliability of this result.

Plots of the overall nominal morphology error rates for each condition and group are shown in Figures 1 and 2. Box plots are utilized due to their effectiveness for visualization of

central tendency and dispersion (Larson-Hall, 2010, p. 245). It is evident from these plots that there was more dispersion of individual error rates within the L2 group than the heritage group in both conditions, while the heritage learners performed more consistently. The plots also demonstrate that overall error rates were relatively similar between the two groups in the timed condition, but considerably less similar in the untimed condition.

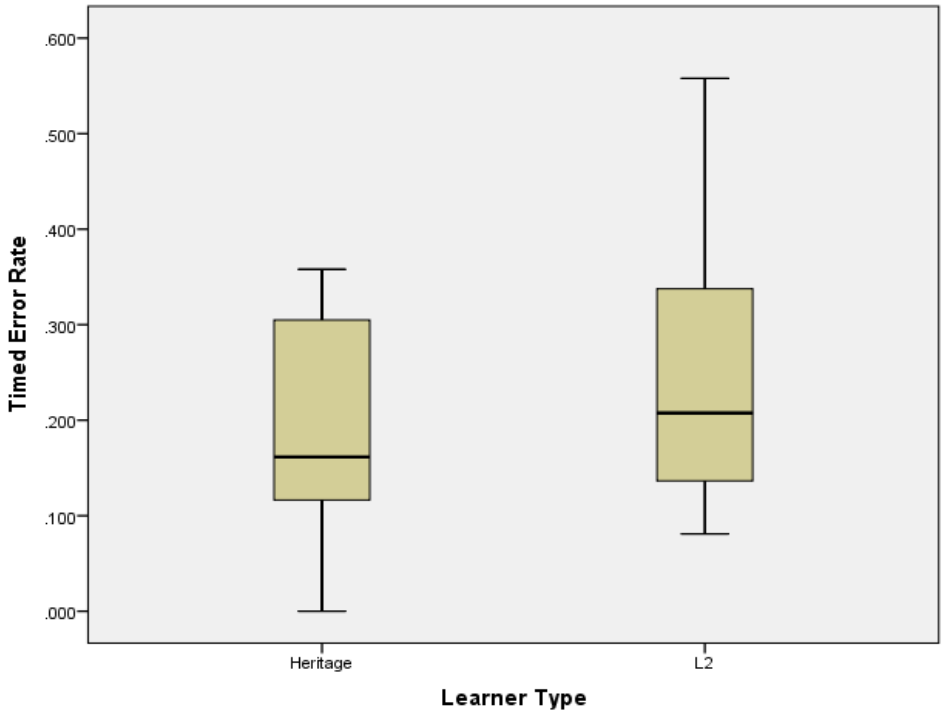

Figure 1: Overall Error Rates per T-unit in Timed Condition by Learner Group

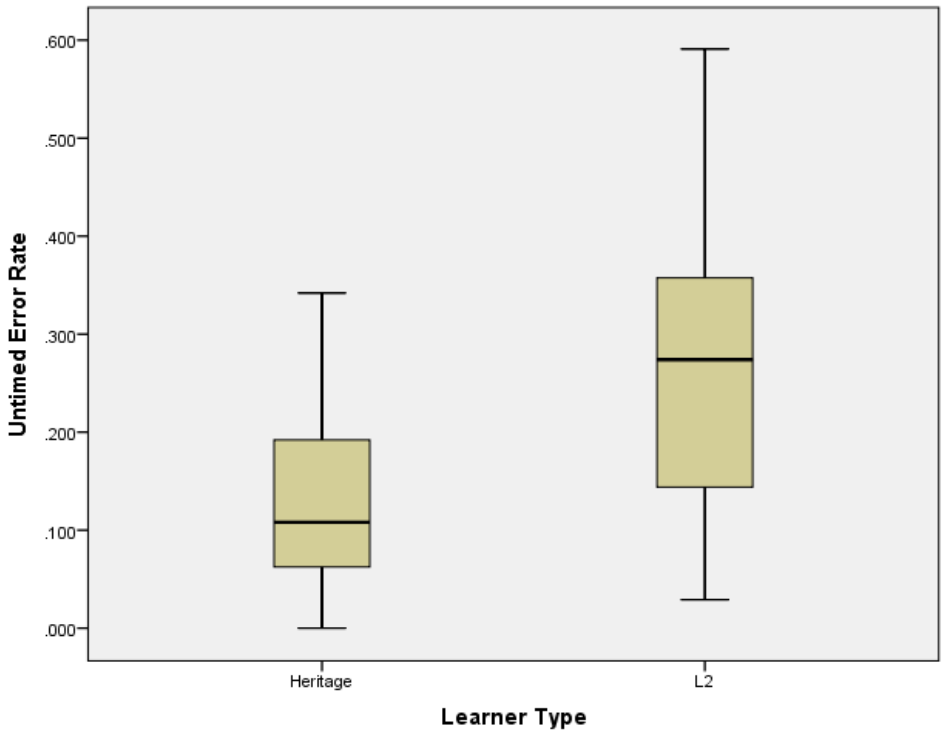

Figure 2: Overall Error Rates per T-unit in Untimed Condition by Learner Group

When rates of case errors and gender errors were analyzed separately, a somewhat more detailed picture emerged, as shown in Table 4.

Table 4: Case and Gender Error Rates per T-unit by Learner Group and Condition

Group/Condition	**Timed Condition**	**Untimed Condition**
Heritage Learner Group:		
Case Error Rates	0.141 (0.108)	0.103 (0.076)
Gender Error Rates *	0.010 (0.012)	0.010 (0.019)
L2 Learner Group:		
Case Error Rates	0.191 (0.116)	0.202 (0.127)
Gender Error Rates *	0.022 (0.024)	0.043 (0.040)

* $p < .05$

There was still no significant main effect of the timed versus untimed condition on either case or gender error rates. However, the main effect of group was nearly significant at the .05 level for

case error rates ($F(1, 22) = 4.236$, $p = .052$) and nearly significant at the .01 level for gender error rates ($F(1, 22) = 7.124$, $p = .014$), with the heritage learners making fewer errors of both types than the L2 learners. Once again, observed power statistics for the main effect of group were relatively low (.503 for case errors and .723 for gender errors), suggesting that reliable group differences might have emerged even more clearly with a larger sample.

Box plots of case and gender error rates for each condition and group are shown in Figures 3-6. It is apparent from these plots that gender error rates were much lower than case error rates for both the heritage and L2 subject groups. The mean rate of gender errors in the heritage group was near zero in both the timed and untimed conditions. Also, both the case and gender error rates within the L2 learner group exhibited greater dispersion than the rates within the heritage group.

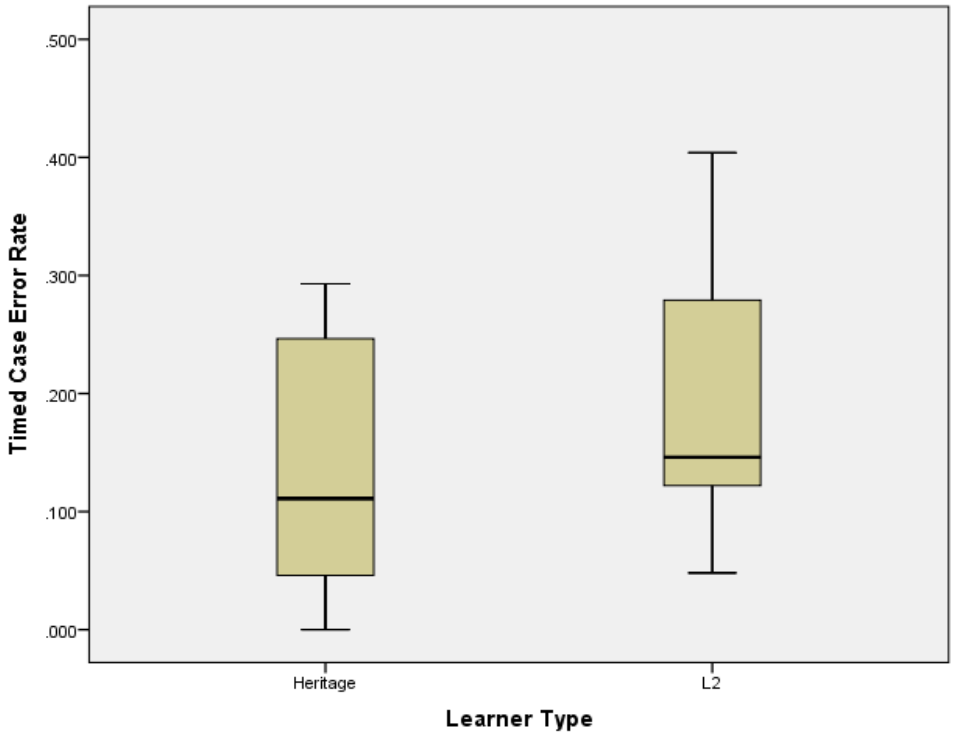

Figure 3: Case Error Rates per T-unit in Timed Condition by Learner Group

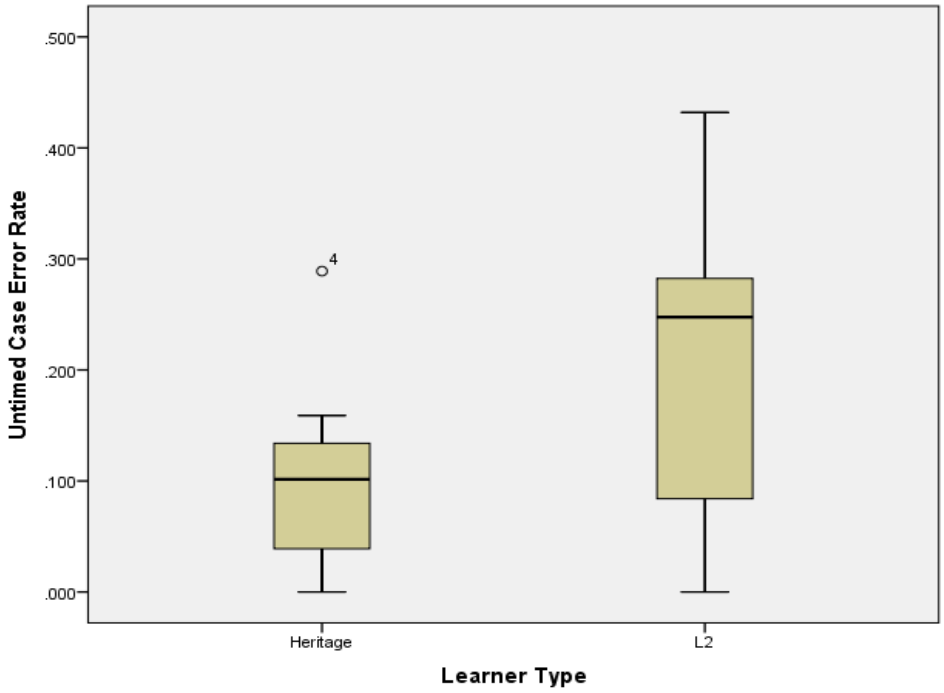

Figure 4: Case Error Rates per T-unit in Untimed Condition by Learner Group

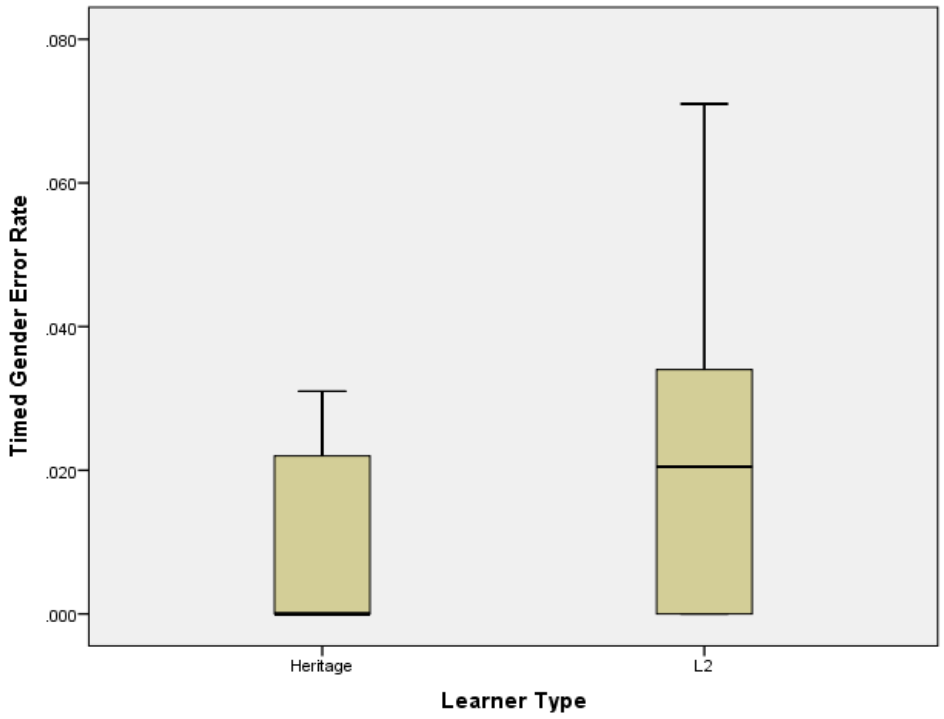

Figure 5: Gender Error Rates per T-unit in Timed Condition by Learner Group

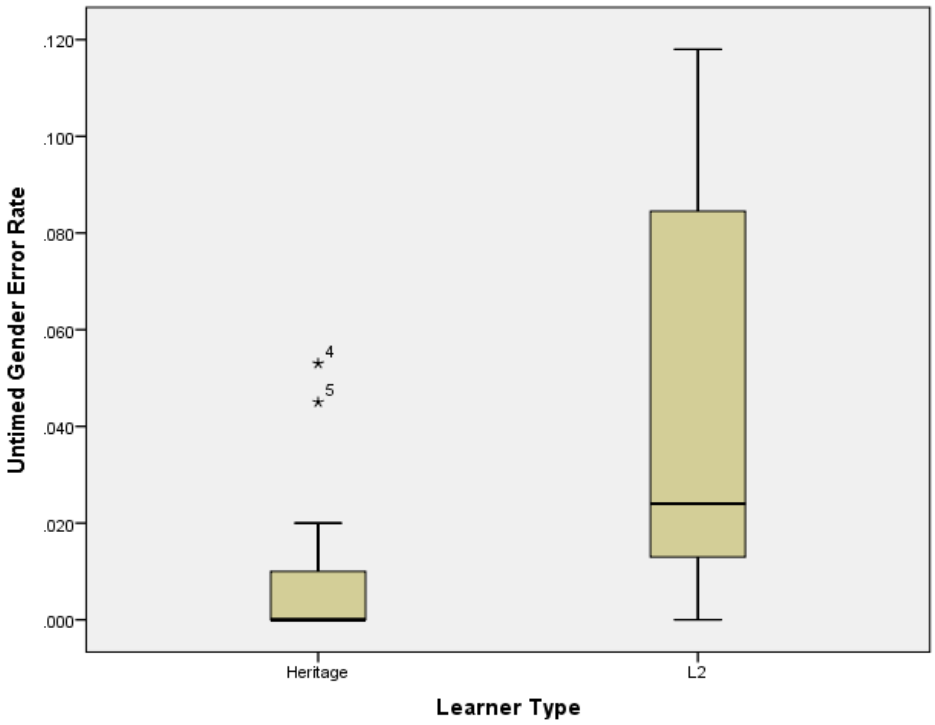

Figure 6: Gender Error Rates per T-unit in Untimed Condition by Learner Group

Returning to the issue of overall error rates, it was suggested above that the unexpected finding of a higher error rate for untimed than timed texts among the L2 learner group could be related to the complexity level of these students' untimed compositions. Not only time pressure, but also texts consisting of longer and more complex phrases and sentences, could contribute to increased production processing difficulties. Thus, in order to investigate whether a relationship existed between text complexity and error rates, an analysis was conducted of the correlation between the difference in complexity levels of timed and untimed texts (measured as words per T-unit) and the difference in overall error rates for timed and untimed texts (measured as errors per T-unit) for each learner within the heritage and L2 subject groups. Differences in both complexity and error rates were expressed as positive numbers when the value for timed texts was greater, and as negative numbers when the value for untimed texts was greater (e.g., if a student's timed texts had an average of 10 words per T-unit, while the same student's untimed

texts had an average of 12 words per T-unit, the difference in complexity for that student would be -2). No statistically significant correlation was found for the heritage learner group. For the L2 learner group, however, the correlation between differences in complexity and error rates for timed and untimed texts approached significance (r = .529, p = .077).

VI. DISCUSSION

The initial prediction that this corpus study would reveal significantly more frequent case and gender errors in the subjects' timed compositions than in their untimed compositions was generally not supported by the data analysis. However, the descriptive statistics for the heritage learner group, with a 45.5% higher overall nominal morphology error rate in the subjects' timed essays, suggest a tendency that should be investigated further with larger samples of heritage learners. While the difference between timed and untimed error rates for the heritage group did not reach statistical significance in this study, it appears to point toward the possibility of errors resulting from processing difficulties.

By contrast, among the L2 learner group, nominal morphology errors actually turned out to be 12.9% more frequent in the students' *untimed* texts than in their timed texts (though this difference also did not reach statistical significance). One possible interpretation of this finding is that the lack of reliable differences between timed and untimed error rates supports a representational deficit approach to explaining the errors made by the advanced Russian language learners in this study, particularly in the L2 group. However, both the L2 and heritage learners had a greater average number of words per T-unit in their untimed essays than in their timed essays. This higher level of complexity of essays written in the untimed condition—when the absence of time limits may have made the students feel capable of attempting more difficult phrase and sentence structures—provides an alternative explanation for the unexpected finding of a higher untimed than timed error rate among the L2 learners.

Yet the correlation between differences in complexity and in error rates for the timed and untimed texts by the L2 group only approached statistical significance; once again, this result would require confirmation in studies with larger numbers of subjects than the 12 L2 and 12 heritage learners included in this study. Error rates for the heritage group, on the other hand, appear to have been affected less by text complexity than by the timed versus untimed status of the students' compositions. The heritage learners had a higher overall error rate in their timed compositions, even though these texts contained fewer words per T-unit than their untimed compositions.

At least in spoken L2 production, increased incidence of inflectional morphology errors in long and complex phrases has been associated with processing difficulties resulting from high cognitive load, as noted in Spinner and Juffs' (2008) study of advanced L1 Italian and Turkish learners of German. This suggests that higher error rates in written production of more complex texts (as with the L2 learner group in the current study), as well as under time pressure (as with the heritage learner group), could potentially be interpreted as consistent with the Missing Surface Inflection Hypothesis (MSIH), since either of these conditions may increase learners' processing difficulties. At the same time, the divergent patterns observed in this study between the two subject groups' susceptibility to different types of processing pressures—combined with the group differences found in their case and gender-marking error rates—indicate intriguing dissimilarities in the processes of language acquisition by Russian heritage versus L2 learners.

The hypothesis that the current study would reveal greater differences in case and gender-marking error rates between timed and untimed compositions by heritage learners than by L2 learners appears to have been supported. In fact, as noted above, not only the magnitude but also the direction of the effect of time pressure on error rates seemed to differ between the heritage

and L2 groups. In addition, there was a statistically significant group difference between the heritage and L2 learners in gender error rates and a nearly significant difference in case error rates, with the heritage learners having lower rates of both types of errors and also displaying less individual variability in error rates than the L2 learners. These findings support the prediction that heritage learners would be less likely than L2 learners to show evidence of possible representational deficits of nominal functional features in their interlanguage grammar.

Also, the fact that case error rates were higher than gender error rates for both groups suggests that acquisition of the Russian case system poses greater difficulties for English-speaking advanced learners than acquisition of the Russian gender system. This may be attributed in part to language-specific factors such as the relatively greater complexity of the case-marking system, since Russian has six grammatical cases compared with three genders, and case but not gender is marked on plural forms. However, this difference in error rates could also be related to the fact that grammatical gender is a lexical feature, while case is primarily a syntactic feature of human languages (though some models of Russian morphosyntax characterize various uses of the oblique cases as lexical). Because some generative theories of language acquisition posit that lexical learning is available throughout the lifespan, while computational learning is fully available only during the critical period, the higher case error rates observed in this study could be viewed as supporting the position that representational problems present a greater obstacle than lexical learning to L2 acquisition.

In summary, due largely to limitations imposed on statistical analysis by the relatively small number of Russian language learners whose essays are included in RULEC, the research questions posed in this paper could not be fully resolved in this exploratory corpus study. However, the descriptive results of this study raise several issues that may stimulate further

research addressing long-running debates on the processing of inflectional morphology, particularly by speakers of a language lacking complex nominal morphology (such as English) who are learning a more morphologically rich language (such as Russian). The differing tendencies found between advanced heritage and L2 learners in this study are especially intriguing in the context of existing literature on heritage language learners.

Previous research on Russian heritage speakers in the U.S. who lack formal instruction in the language has revealed systematic patterns of reduction and reanalysis of both the Russian case system (Polinsky, 1996) and gender system (Polinsky, 2008) in their grammar. Yet the current study demonstrated that Russian heritage learners pursuing advanced levels of instruction in the language could achieve a higher degree of case and gender-marking accuracy in written production than their traditional L2 learner classmates, while also appearing better able to maintain this accuracy in more complex texts. This suggests that heritage learners' exposure to Russian language input during childhood gives them some advantage in acquiring Russian inflectional morphology through formal instruction in adulthood, despite initial systematic gaps in their heritage language grammar and other recent findings that "heritage re-learners do not perform better than L2 learners in morphosyntax" (Polinsky, 2015, p. 169). Polinsky cited several examples involving both nominal and verbal constructions to illustrate this point. Larger-scale studies of advanced Russian heritage and L2 learners could help to reconcile these apparently conflicting accounts.

Other questions that could be productively explored in subsequent studies include whether advanced learners of Russian exhibit higher case and gender-marking error rates in nominal phrases containing relatively infrequent inflectional patterns or lexical items (to investigate possible type and token frequency effects), as well as in nominal phrases that include

adjectives and thus may be more difficult to process than shorter phrases consisting of nouns only. Such fine-grained analysis of learner corpus data could yield additional useful insights into the nature of morphological processing by second language learners whose L1 has a less complex inflectional system.

Finally, the findings of this exploratory study that processing difficulties seemed to affect the incidence of case and gender-marking errors among advanced English-speaking learners of Russian—particularly in the heritage group—suggest that explicit instruction and conscious self-monitoring strategies may be effective in reducing morphological errors among such students. Representational deficit models of L2 morphological processing hold that instruction only enhances adult learners' declarative knowledge, but cannot develop the procedural or implicit knowledge that is required for native-like processing; while connectionist models indicate that explicit instruction may be a less effective method than input flood for improving the accuracy of learners' form-meaning mapping in inflectional morphology. However, the conclusions that time pressure (for heritage learners) and text complexity (for traditional L2 learners) appear to increase error rates do not provide support for either of those types of models. If future research utilizing larger subject groups confirms these findings, this development would have important implications for foreign language programs at U.S. universities that seek to facilitate attainment of professional-level proficiency by learners of Russian and other critical languages.

APPENDIX A

XML ANNOTATION SYSTEM USED FOR ERROR TAGGING

Basic format of error tags:

```
<err cf = 'correct form of word' pos = 'nn'|'adj'|'det'
gnd = 'ms'|'fm'|'nt' num = 'sg'|'pl' cs = 'n'|'a'|'g'|'d'|'i'|'p'
t = 'gnd'|'num'|'cs'|'un'>incorrect form of word</err>
```

Possible attribute values for part of speech; gender, number and case of the correct form of the tagged word (not the incorrect form used by the text author); and error type:

Part of speech (pos)
nn = noun (including proper nouns; excluding pronouns)
adj = adjective
det = determiner (including equivalents of “who”, “which”)

Gender (gnd)
ms = masculine
fm = feminine
nt = neuter

Number (num)
sg = singular
pl = plural

Case (cs)
n = nominative
a = accusative
g = genitive
d = dative
i = instrumental
p = prepositional

Type of error (t) (for this attribute only, multiple values may be chosen, but ‘un’ may not be combined with any other value)
gnd = gender
num = number
cs = case
un = unclear

APPENDIX B

EXAMPLE OF ANNOTATED TEXT FROM RULEC CORPUS

<essay>Если бы я снимала короткий фильм о славянском народе, я бы выбрала русский народ, потому что он мне ближе и более знаком. В этом фильме я бы рассказала о древних традициях, о древних богах, идолах. О их вере. Так же вставила бы изображения домов и природы. Я думаю, что <err cf="моей" pos="det" gnd="fm" num="sg" cs="i" t="cs">моя</err> <err cf="главной" pos="adj" gnd="fm" num="sg" cs="i" t="cs">главная</err> <err cf="темой" pos="nn" gnd="fm" num="sg" cs="i" t="cs">тема</err> все же были бы их традиции. Потому что, насколько я помню, у них было очень много богов и они поклонялись идолам.</essay>

APPENDIX C

PYTHON SCRIPT USED TO COUNT ERRORS BY CATEGORY

```
import glob, bs4
rulecfiles = glob.glob('C:/Users/Gina Peirce/Documents/Linguistics/
    Thesis Research/RULEC/ANNOTATED_XML/*')
errortotal = 0
casetotal = 0
gendertotal = 0
numbertotal = 0
csgndtotal = 0
csnumtotal = 0
gndnumtotal = 0
uncleartotal = 0
# sorts essay files into four groups and goes through one group at a
    time
conditionals = ['_HL_,non-timed', '_HL_,_timed', '_FL_,non-timed',
    '_FL_,_timed']
x = 0
while x < 4:
    print '\n\n', conditionals[x], '\n'
    cond = conditionals[x].split(',')
    fileindex = 0
    for t in rulecfiles:
        if cond[0] in t and cond[1] in t:
            f = open(t)
            text = f.read()
            f.close()
            soup = bs4.BeautifulSoup(text)
            # counts errors in each individual essay using
               BeautifulSoup 4 package
            errs = soup.find_all ('err')
            cserrs = soup.find_all('err', {'t':'cs'})
            gnderrs = soup.find_all('err', {'t':'gnd'})
            numerrs = soup.find_all('err', {'t':'num'})
            csgnderrs = soup.find_all('err', {'t':'cs gnd'})
            gndcserrs = soup.find_all('err', {'t':'gnd cs'})
            csnumerrs = soup.find_all('err', {'t':'cs num'})
            numcserrs = soup.find_all('err', {'t':'num cs'})
            gndnumerrs = soup.find_all('err', {'t':'gnd num'})
            numgnderrs = soup.find_all('err', {'t':'num gnd'})
```

```
            unerrs = soup.find_all('err', {'t':'un'})
            # prints error totals for each individual essay
            print rulecfiles[fileindex][79:]+':\n\t',
            print  len(cserrs)  +  len(csgnderrs)  +  len(csnumerrs)  +
               len(gndcserrs) + len(numcserrs), 'case error(s),',
            print  len(gnderrs)  +  len(csgnderrs)  +  len(gndnumerrs)  +
               len(gndcserrs) + len(numgnderrs), 'gender error(s),',
            print len(errs), 'total error(s)'
            # adds these to error totals for entire sample
            errortotal = errortotal + len(errs)
            csgndtotal = csgndtotal + len(csgnderrs) + len(gndcserrs)
            csnumtotal = csnumtotal + len(csnumerrs) + len(numcserrs)
            gndnumtotal    =    gndnumtotal    +    len(gndnumerrs)    +
               len(numgnderrs)
            casetotal  =  casetotal  +  len(cserrs)  +  len(csgnderrs)  +
               len(csnumerrs) + len(gndcserrs) + len(numcserrs)
            gendertotal = gendertotal + len(gnderrs) + len(csgnderrs)
               + len(gndnumerrs) + len(gndcserrs) + len(numgnderrs)
            numbertotal = numbertotal + len(numerrs) + len(csnumerrs)
               + len(gndnumerrs) + len(numcserrs) + len(numgnderrs)
            uncleartotal = uncleartotal + len(unerrs)
        fileindex +=1
    fileindex = 0
    x +=1
print  '\n\n',  'The  entire  sample  has',  errortotal,  'total  errors,
    including:'
print casetotal, 'case errors'
print gendertotal, 'gender errors'
print numbertotal, 'number errors'
print 'and', uncleartotal, 'errors of unclear type.'
```

BIBLIOGRAPHY

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language, 40*, 41-61.

Alsufieva (Yatsenko), A., Kisselev, O., & Freels, S. (2012). Results 2012: Using Flagship data to develop a Russian Learner Corpus of Academic Writing. *Russian Language Journal 62*, 79-105.

Clahsen, H. (1995). German plurals in adult second language development: Evidence for a dual-mechanism model of inflection. In Eubank, L., Sharwood-Smith, M., & Selinker, L. (eds.) *The current state of interlanguage: Studies in Honour of William Rutherford*. Philadelphia: John Benjamins. 123-138.

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning, 60*(1), 21-43.

Dabrowska, E. (2008). The effects of frequency and neighborhood density of adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language, 58*, 931-951.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.

Gor, K. (2010a). Introduction. Beyond the obvious: Do second language learners process inflectional morphology? *Language Learning, 60*(1), 1-20.

Gor, K., & Cook, S. (2010b). Nonnative processing of verbal morphology: In search of regularity. *Language Learning, 60*(1), 88-126.

Gor, K., & Vdovina, T. (2010c). Frequency, regularity and input in second language processing of Russian verb inflection. *The Slavic and East European Journal, 54*(1), 7-31.

Hawkins, R., & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis.' *Second Language Research, 13*, 187-226.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*(1), 33-56.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603-634.

Kazanina, N., Dukova-Zheleva, G., Geber, D., Kharlamov, V., & Tonciulescu, K. (2008). Decomposition into multiple morphemes during lexical access: A masked priming study of Russian nouns. *Language and Cognitive Processes, 23*(6), 800-823.

Kempe, V., & MacWhinney, B. (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition, 20*, 543-587.

Krashen, S. (2009). *Principles and Practice in Second Language Acquisition* (First Internet Edition). Retrieved from http://www.sdkrashen.com/content/books/principles_and_practice.pdf

"Language Flagship: Creating global professionals." (2013). The Language Flagship. Retrieved from http://www.thelanguageflagship.org/

Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York: Routledge.

McDonald, J. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*, 381-401.

Polinsky, M. (1996). American Russian: An endangered language? (Manuscript). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.138.1625&rep=rep1&type=pdf

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal, 6*(1), 40-71.

Polinsky, M. (2015). When L1 becomes an L3: Do heritage speakers make better L3 learners? *Bilingualism: Language and Cognition, 18*, 163-178.

Portin, M., Lehtonen, M., Harrer, G., Wande, E., Niemi, J., & Laine, M. (2008). L1 effects on the processing of inflected nouns in L2. *Acta Psychologica, 128*, 452-465.

Portin, M., Lehtonen, M., & Laine, M. (2007). Processing of inflected nouns in late bilinguals. *Applied Psycholinguistics, 28*, 135-156.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition. *Second Language Research, 16*(2), 103-133.

Rehak, K., & Juffs, A. (2011). Native and non-native processing of morphologically complex English words: Testing the influence of derivational prefixes. In G. Granena (ed.) *Selected Proceedings of the 2010 Second Language Acquisition Research Forum*. College Park, MD: Cascadilla Press. 125-142.

"RULEC: Russian Learner Corpus of Academic Writing." (2009). Corpus Technologies. Retrieved from http://web-corpora.net/RussianLearnerCorpus/search/

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.

Spinner, P. (2013). The second language acquisition of number and gender in Swahili: A Feature Reassembly approach. *Second Language Research 29* (4), 455-479.

Spinner, P., & Juffs, A. (2008). L2 grammatical gender in a complex morphological system: The case of German. *International Review of Applied Linguistics 46*, 315-348.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. New York: Springer.

Taraban, R., & Kempe, V. (1999). Gender processing in native and nonnative Russian speakers. *Applied Psycholinguistics, 20*, 119-148.

Tkachenko, E. & Chernigovskaya, T. (2010). Input frequencies in processing of verbal morphology in L1 and L2: Evidence from Russian. In Gronn, A., & Marijanovic, I. (eds.) *Russian in Contrast: Lexicon. Oslo Studies in Language 2*(2), 281-318.