

Representational and Processing Constraints on the Acquisition of Case and Gender by Heritage and L2 Learners of Russian: A Corpus Study

Gina Peirce
University of Pittsburgh

ABSTRACT

Morphological errors are prevalent in adult second language production, particularly among learners whose first languages have less complex inflectional systems. Thus, U.S. learners of Russian can provide a testing ground for competing approaches to L2 morphological acquisition. This study utilizes the Russian Learner Corpus of Academic Writing (2017) to compare case and gender-marking error frequencies in timed versus untimed essays by advanced heritage and traditional L2 learners in Portland State University's Russian Language Flagship program.

It was predicted that higher error rates in timed compositions would support the position that advanced learners' morphological errors reflect processing difficulties under time pressure. However, such differences did not reach significance for either heritage or L2 learners; in the latter group, error rates were higher in students' *untimed* texts. These results could be interpreted as demonstrating representational deficits in interlanguage grammar, particularly in the L2 group. However, greater complexity (words per T-unit) of the untimed essays provides an alternative explanation for the higher untimed error rate among this group. The heritage group had lower overall case and gender-marking error rates than the L2 group, suggesting heritage learners are less likely to show evidence of possible representational deficits of nominal functional features in their interlanguage grammar.

Keywords: *Russian, morphology, case, gender, processing, corpus*

INTRODUCTION

This study addresses the prevalence of inflectional morphology errors in the production of adult language learners, including learners with a high overall proficiency level and many years of exposure to the L2—a phenomenon for which second language acquisition researchers have proposed various explanations. These include such diverse causes as L1 influence (e.g., Portin, Lehtonen, & Laine, 2007; Portin et al., 2008); online processing difficulties resulting from high cognitive load (e.g., Prévost & White, 2000; McDonald, 2006); the weakness of morphosyntactic cues to form-meaning mapping (e.g., Kempe & MacWhinney, 1998; Taraban & Kempe, 1999); and a “bottleneck” in L2 acquisition that results from the encoding of all formal features of the grammar in inflectional morphology (e.g., Slabakova, 2009). 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 & Chan, 1997), or to adult learners' lack of access to language-specific implicit learning mechanisms (e.g., Clahsen, Felser, Neubauer, Sato, & Silva, 2010). Under the representational deficit approach, learners' mental representations of L2 structures in their interlanguage grammar inevitably remain divergent from the mental representations held by native speakers of the target language.

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. Comparative examination of U.S. heritage and traditional second language learners acquiring the Russian case and gender systems can provide a testing ground for competing approaches to L2 morphological processing, as well as exploring the possibility that heritage learners' early childhood exposure to Russian language input results in different patterns of acquisition of nominal morphology than are found among non-heritage L2 learners.

This article begins by reviewing studies of acquisition of inflectional morphology, focusing on the following questions: (1) Are errors in case and gender marking by English-speaking 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? (3) Do heritage and traditional L2 learners exhibit different patterns and outcomes of acquisition of Russian nominal morphology?

The article then presents an analysis of the Russian Learner Corpus of Academic Writing (RULEC, 2017), a collection of texts written by advanced heritage and L2 learners in the Russian Flagship program at Portland State University. This analysis compares 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. In other words, learners may experience difficulty with basic cognitive processes such as working memory, decoding, and processing speed when engaging in real-time, or online, language production and comprehension tasks that place high demands on their cognitive resources. Such processing difficulties may interfere with learners' ability to access and apply their existing grammatical knowledge to these types of tasks (e.g., McDonald, 2006, p. 382). Conversely, a lack of significant differences in error rates in the timed versus untimed compositions in RULEC 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, including findings of divergent patterns among the heritage and L2 learner groups; these are detailed in the discussion section of this paper.

THEORETICAL FRAMEWORK AND PREVIOUS RESEARCH

Models of Morphological Processing and Acquisition

The research questions addressed here are situated in a framework of competing theoretical models of processing and acquisition of inflectional morphology, which have varying implications for the possible causes of language learner errors. As outlined in Gor (2010), dual-system approaches make a categorical distinction between rule-based learning of regular inflections and associational learning of irregular inflections; while single-system approaches 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. Alternatively, a usage-based approach (e.g., 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 support a usage-based account, finding that L2 acquisition of Russian verbal inflection is affected by both type and token frequency.

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, 2010, p. 5). Thus, Gor concluded that a categorical distinction cannot be universal and proposed a continuum between regular and irregular processing, suggesting that 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 (2010) study of production and recognition of conjugated verbs by Russian L2 and heritage learners. 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 L2 learners.

Dual-system approaches to morphological processing can accompany theories proposing fundamental distinctions between L1 and L2 acquisition; for instance, Clahsen et al. (2010) posited representational deficits in the inflectional grammar of adult L2 learners. Clahsen 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, 2010, p. 10). This view was 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).

A review by Clahsen et al. of online processing studies concluded that even advanced L2 learners rely more on declarative memory than decomposition to process morphologically complex words. This position is consistent with Ullman’s theory that the procedural or implicit knowledge system is attenuated in adult language learners (Clahsen, 2010, p. 38). The declarative/procedural model of language acquisition contrasts declarative memory—including lexical learning, as well as explicit or conscious knowledge of the grammatical system of a language—with procedural memory, which includes implicit or unconscious knowledge of grammar that some researchers claim can be fully acquired only by L1 speakers and those learning an L2 in childhood (e.g., Ullman, 2013). Morphological decomposition is based on procedural memory, while storage of full-form representations of inflected words depends on declarative memory.

Jiang's (2004) account of representational deficits in the grammar of L2 learners pointed also to the role of the particular L1, demonstrating through a self-paced reading study that native Chinese-speaking learners of English were insensitive to morphological plural markers and number agreement in their L2. Jiang attributed this finding to a deficiency in the learners' implicit mental representations that could result from age-related critical period effects combined with the absence of plural morphology in their Chinese L1 (p. 627), which required them to store English number agreement only within their explicit or declarative knowledge.

Other researchers have presented opposing views on L2 morphological acquisition, such as Prévost and White's (2000) study testing the Missing Surface Inflection Hypothesis (MSIH) against the Impaired Representation Hypothesis (IRH). The MSIH posited that learners have unconscious knowledge of L2 morphological features, but experience difficulty with realization of surface forms in online processing, while the IRH suggested permanent impairment of interlanguage grammar at the level of functional categories or features in learners' mental representations (Prévost & White, 2000, pp. 108-110). A related account to the IRH was the Failed Functional Features Hypothesis or FFFH, proposed by Hawkins and Chan (1997), which held 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. These findings supported the MSIH, since the IRH would have predicted arbitrary use of inflectional morphology with finite and non-finite forms being used randomly.

In a study of gender agreement errors by two L2 German learners (one L1 Italian and one L1 Turkish), Spinner and Juffs (2008) also reported evidence of unidirectional errors in their oral production, which—as in Prévost and White (2000)—indicated systematicity rather than randomness in subjects' interlanguage and thus was consistent with the MSIH. In addition, Spinner and Juffs found that the subjects' errors were not attributable to the absence of relevant functional features in their respective native languages, which contradicted the FFFH. Instead, these learners' frequent omission or defaulting of inflected forms under conditions of high cognitive load, such as long or complex phrases, pointed toward processing difficulties.

McDonald (2006) also focused on cognitive load to demonstrate that processing difficulties were a likelier cause of L2 learner errors than representational deficits. 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, Ayres, & Kalyuga, 2011, pp. 52-53), such as adult second language acquisition. 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. McDonald found that L2 learners of English in an unstressed condition performed similarly on grammaticality judgment tests 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. (2007 and 2008) provided further evidence that adult L2 learners could develop mental representations similar to those of native speakers—contrary to the representational deficit approach—while also highlighting the role of L1 effects on acquisition of L2 morphology. Portin et al. (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. The authors concluded that “lengthy formal study [...] can provide a late learner with nativelike full-form input representations” (p. 151). This finding contradicted Clahsen’s (2010) position that L2 learners rely more than native speakers on full-form representations rather than decomposition of inflected words regardless of frequency, suggesting instead that with continued exposure to target language input, learners’ mental representations become more native-like over time.

In another L2 lexical decision study, Portin et al. (2008) found that L1 Hungarian and Chinese speakers transferred strategies from their respective native languages when processing inflected Swedish nouns. Despite claims that L1 effects are absent from L2 processing, e.g. by Silva and Clahsen (cf. Rehak & Juffs, 2011, p. 126), Portin et al. presented compelling evidence that the structures of a learner’s particular L1 and L2 must be considered. As indicated in Gor (2010), the relative importance of decomposition and full-form storage may vary with morphological richness of specific languages. Processing of full-form representations is more achievable 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.

Finally, others have argued for the possibility of native-like representations among L2 learners, and therefore against representational deficits, 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 would 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). 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).

Heritage Language Morphology

Heritage speakers, who do not fit neatly into an L1/L2 dichotomy, have become an important topic of morphological processing research. For instance, Montrul, Foote, & Perpiñan (2008) found that the presence of systematic gender agreement errors on experimental tasks among Spanish heritage speakers studying the language at college level contradicted representational deficit accounts of morphological acquisition. Given their early childhood exposure to Spanish, these learners should not have exhibited such errors if they were caused by critical period effects (p. 539). On the other hand, the heritage learners outperformed L2 Spanish learners at an equivalent proficiency level on an oral production task while the L2 group made fewer errors on written tests of comprehension of gender agreement, suggesting a possible advantage for heritage speakers on tasks measuring implicit linguistic knowledge (p. 541). Montrul, Davidson, de la Fuente, & Foote (2014) found that Spanish heritage learners processed gender morphology similarly to monolingual native speakers on an oral task targeting implicit knowledge, yet did not differ significantly from L2 learners on more explicit grammatical judgment tasks (p. 134). These results indicate that the context of initial language exposure (school-based instruction emphasizing explicit knowledge, versus a naturalistic home environment) may affect outcomes of morphological acquisition.

Polinsky's (2008) experiments with Russian heritage speakers who lacked formal instruction in the language also 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). Polinsky (1996) demonstrated that other components of the Russian grammatical system, including case morphology—as reflected in oral production by U.S. heritage speakers who did not complete primary schooling in Russian—exhibited the effects of systematic reduction and reanalysis as well (p. 27).

Thus Polinsky, like Montrul, 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" (Polinsky, 2008, p. 64). This paper 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, this literature review shows that while some morphological processing studies have found evidence for representational deficits in adult language learners through online tasks such as self-paced reading (e.g., Jiang, 2004) or masked priming (e.g., Silva & Clahsen, 2008, as summarized in Rehak & Juffs, 2011), several others have presented persuasive evidence that inflectional morphology errors may be due to other causes, including the increased processing load associated with timed versus untimed tasks. These findings have resulted from studies utilizing

both oral production data (e.g., Prévost & White, 2000; Spinner & Juffs, 2008; and Polinsky, 2008) and input processing tasks such as lexical decision (e.g., Portin et al., 2007, 2008), grammaticality judgment (e.g., McDonald, 2006), and self-paced reading (e.g., Taraban & Kempe, 1999). This array of evidence suggests that representational deficit approaches have not adequately accounted for language learner difficulties with inflectional morphology.

Additionally, the review indicates that the relatively few existing studies on processing of Russian case and gender by L1 English adult learners (e.g., Kempe & MacWhinney, 1998; Taraban & Kempe, 1999) have examined this phenomenon mainly from a connectionist perspective. There is a need 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 paper contributes to addressing these needs, employing the recommendation from Polinsky (2008) to compare heritage speakers with traditional L2 learners. The corpus study detailed below investigates 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 also explores whether different patterns can be observed as part of the effects of time pressure on morphological accuracy of written production by Russian heritage versus non-heritage learners.

METHODOLOGY

Data Source

All data for this study were obtained from RULEC (2017), the Russian Learner Corpus of Academic Writing. The use of learner corpora is a somewhat recent development in second language acquisition 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 to address “the relative lack of empirical research on advanced interlanguage” (Alsufieva, Kisselev, & Freels, 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 use of this learner corpus affords an opportunity 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. Thus, a corpus study is an appropriate methodology for investigating the questions about L2 difficulties with inflectional morphology raised in this paper. 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 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” (Alsufieva et al., 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, 2017). 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 (The Language Flagship, 2013).

All students had been administered Oral Proficiency Interviews in Russian and had been rated between Intermediate Mid and Advanced High on the American Council on the Teaching of Foreign Languages (ACTFL) scale (ACTFL, 2012) 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 in this group than in traditional L2 learners, as suggested by Polinsky (2008) and Gor & Cook (2010).

Each text in the corpus is tagged according to criteria including the L2 or heritage status and oral proficiency level of the author, whether the writing assignment was timed or untimed, 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 2,250 T-units.² 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 made available to researchers included 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 equitable representation of the remaining authors in RULEC (O. Kisselev, personal communication, October 21, 2014). It was not feasible to increase the number of subjects

included in this study from 12 to 13 each in the heritage and L2 groups, due to an insufficient number of students with enough timed texts in RULEC that were comparable in length to the same individuals' untimed texts.

The limited number of subjects available through RULEC also prevented precise matching of the heritage and L2 groups by oral proficiency level. However, differences between the two subject groups were relatively small. Among the 116 texts per group used in the study sample, over 78% in each group were authored by students whose proficiency was rated either Intermediate High or Advanced Low at the time when they wrote the essays, with the largest number of essays in both groups produced by students at the Advanced Low level. The sample texts in each group also included a small number written by students at the Intermediate Mid and Advanced Mid proficiency levels; no texts by students at the Advanced High level were included in the study sample. These statistics are based on individual texts rather than on overall ratings for each learner, since most of the students' measured proficiency levels changed at least once over the period when their essays were being collected for the corpus.

The heritage group did have a somewhat higher number of texts by students at the Advanced Low versus Intermediate High level relative to the L2 group: 71 Advanced Low and 20 Intermediate High in the heritage group, compared to 54 Advanced Low and 46 Intermediate High in the L2 group. Yet it should also be noted that fluency—as demonstrated by the ability to sustain paragraph-length rather than sentence-length discourse—plays a greater role than morphological accuracy in enabling learners to transition from the Intermediate High to the Advanced Low level on the ACTFL oral proficiency scale (ACTFL, 2012). Thus, a rating of Advanced Low proficiency does not always indicate a much higher degree of grammatical accuracy than a score of Intermediate High.

Data Preparation and Analysis

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.³ The entire dataset of 232 annotated texts, along with an explanation of the annotation system used for error tagging, is publicly accessible through the University of Pittsburgh's digital scholarship repository (Peirce, 2016).

Because Russian is a synthetic language with a single inflectional ending to mark case, gender and number, and because of 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 for 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% of errors, were tagged as being of an unclear type.

To verify the accuracy of error annotation, the author of this paper 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 an item constituted a nominal morphology error were present for six tags, while differences in categorization of the type of error (e.g., case error, gender error, or morphological error of unclear type) were present for another nine 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. In linguistic studies of annotated corpora, agreement rates that are considered acceptable vary according to a study's purposes (e.g., Artstein & Poesio, 2008, p. 557); but 100% inter-annotator agreement is not typically attainable. In the current study, the need for the annotator to interpret the meanings intended by essay authors in instances of possible morphological errors likely accounted for some differences in detection of the presence and/or type of errors.

After errors were tagged in the selected texts, 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; all T-units in each of the selected texts were annotated with XML tags by the author of this paper.

In addition to overall error counts, this study used a Python script 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 the use of an allomorph associated with "hard" adjectives where an allomorph associated with "soft" adjectives (those with stems ending in palatalized [n^j]) 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.

PREDICTIONS

The methodology for this study utilized a corpus of written production data by language 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.

First, 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,⁴ 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.

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 were 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.

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	529
Heritage Group – Untimed	74	522
L2 Group – Timed	146	626
L2 Group – Untimed	168	573

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

Group/Condition	Timed Condition	Untimed Condition
Heritage Learner Group *	0.187 (0.113)	0.133 (0.104)
L2 Learner Group *	0.247 (0.158)	0.282 (0.166)

* $p < .05$ (main effect of group)

Analysis of the sample of RULEC texts annotated for this study revealed that the heritage learner group had an overall rate of 0.133 nominal morphology errors per T-unit in their untimed compositions, while their overall error rate in timed compositions was 40.6% higher at 0.187 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 14.2% higher in their untimed compositions (0.282 errors per T-unit) than in their timed compositions (0.247 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 are reported below.

The repeated measures ANOVA analysis did not reveal a significant main effect of the within-subjects variable of timed versus untimed condition ($F(1, 22) = .099, p = .756$) or a significant effect of the interaction of condition and the between-subjects variable of learner group ($F(1, 22) = 2.134, p = .158$), on overall error rates. However, the main effect of group was significant at the .05 level ($F(1, 22) = 4.810, p = .039$), with a higher error rate for the L2 learner group (0.264 per T-unit) than for the heritage learner group (0.160 per T-unit). It should nonetheless be noted that the observed power statistic of .554 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.

It is evident from the standard deviations reported in Table 3 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. Overall error rates were relatively similar between the two groups in the timed condition, but considerably less similar in the untimed condition.

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 *	0.136 (0.101)	0.107 (0.081)
Gender Error Rates *	0.009 (0.011)	0.010 (0.020)
L2 Learner Group:		
Case Error Rates *	0.189 (0.116)	0.204 (0.128)
Gender Error Rates *	0.021 (0.024)	0.043 (0.040)

* $p < .05$ (main effect of group)

There was still no significant main effect of the timed versus untimed condition on either case or gender error rates, although the main effect of condition did approach significance for gender errors only ($F(1, 22) = 3.423, p = .078$) with a higher error rate in the untimed condition, primarily among the L2 learners. However, the main effect of group was significant at the .05 level for case error rates ($F(1, 22) = 4.385, p = .048$) and nearly significant at the .01 level for gender error rates ($F(1, 22) = 7.161, p = .014$), 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 (.517 for case errors and .725 for gender errors), suggesting that reliable group differences might have emerged even more clearly with a larger sample.

It is apparent 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.

Returning to the issue of overall error rates, it was suggested above that the unexpected finding of a higher error rate for untimed than for timed texts in 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 ($r = .499, p = .099$).

DISCUSSION AND FUTURE DIRECTIONS

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 40.6% 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 14.2% 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 L2 group, such as the Impaired Representation Hypothesis (IRH; see discussion in Prevost & White, 2000) or the Failed Functional Features Hypothesis (FFFH; see Hawkins & Chan, 1997). However, both the L2 and heritage learners in this study 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's (2008) study of advanced L2 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), and of texts being composed 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, statistically significant group differences were found between the heritage and L2 learners in overall, case-marking, and gender-marking error rates, with the heritage learners having lower rates of each type of error 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.

Moreover, the consistency for both groups of higher case error rates than gender error rates 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. Both Gor (2010) and Portin et al. (2008) have proposed that processing strategies vary with the morphological richness of particular grammatical systems. Additionally, the subjects of this study might to some extent have avoided using nouns of whose gender they were unsure. However, the difference in error rates could also reflect the differences between grammatical gender, which is a lexical feature, and case, which 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 at a younger age, 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 the 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 its gender system (Polinsky, 2008) in their grammar, and morphological errors have been found to persist among learners studying other heritage languages, such as Spanish at the college level (e.g., Montrul et al., 2008). The current study has demonstrated that Russian heritage learners pursuing advanced levels of instruction 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 may give

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 provides 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 and clarify whether the different findings primarily reflect variation among populations of heritage speakers in diverse settings.

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 potential type and token frequency effects posited in usage-based accounts such as Dabrowska, 2008), 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 decreasing morphological errors among such students. Representational deficit models of L2 morphological processing (e.g., Clahsen et al., 2010; Ullman, 2013) hold that instruction only enhances adult learners’ declarative knowledge, but cannot develop the procedural or implicit knowledge that is required for native-like processing. Connectionist models (e.g., Kempe and MacWhinney, 1998) indicate on the other hand 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, this study’s conclusions that time pressure for heritage learners, and text complexity for traditional L2 learners, appears to increase error rates do not provide support for either of those types of models. Instead, these findings suggest that morphological errors often stem from processing difficulties which may be reduced over time through instructed learning, thus promoting the eventual emergence of more native-like processing and production abilities among advanced learners of Russian.

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 morphologically complex languages. The results of this study suggest that heritage learners have especially strong potential to achieve high levels of accuracy in language production through a well-designed instructional program, including in the area of inflectional morphology, which is problematic even for advanced second language learners.

REFERENCES

- 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.
- American Council on the Teaching of Foreign Languages (ACTFL). (2012). ACTFL Proficiency Guidelines 2012. Retrieved from https://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf
- Artstein, R., & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4), 555-596.
- 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 Honor of William Rutherford*. Philadelphia, PA: 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 on 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, UK: Oxford University Press.
- Gor, K. (2010). Introduction. Beyond the obvious: Do second language learners process inflectional morphology? *Language Learning*, 60(1), 1-20.
- Gor, K., & Cook, S. (2010). Nonnative processing of verbal morphology: In search of regularity. *Language Learning*, 60(1), 88-126.
- 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.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford, UK: Oxford University Press.
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, 25, 603-634.
- 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.

- Kisselev, O., & Alsufieva, A. (compilers). RULEC RLC Subcorpus. (2017). School of Linguistics National Research University Higher School of Economics. Retrieved from <http://www.web-corpora.net/RLC/rulec>
- 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
- 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.
- Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58(3), 503-553.
- Montrul, S., Davidson, J., de la Fuente, I., & Foote, R. (2014). Early language experience facilitates the processing of gender agreement in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 17(1), 118-138.
- Oxygen XML Editor (19.1). [Data software]. (2017). Craiova, Romania: Syncro Soft.
- Peirce, G. (2016). Annotated sample of timed and untimed essays by heritage and second language learners from the Russian Learner Corpus of Academic Writing (RULEC). [Dataset]. University of Pittsburgh. Retrieved from <http://d-scholarship.pitt.edu/29214>
- Polinsky, M. (1996). American Russian: An endangered language? [Manuscript]. University of California, San Diego. 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.
- Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, 11(2), 245-260.
- Slabakova, R. (2009). What is easy and what is hard to acquire in a second language? In Melissa Bowles et al. (Eds.) *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)* (pp. 280-294). Somerville, MA: Cascadilla Proceedings Project.
- 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, NY: Springer.
- Taraban, R., & Kempe, V. (1999). Gender processing in native and nonnative Russian speakers. *Applied Psycholinguistics*, 20, 119-148.
- The Language Flagship. (2013). The Language Flagship: Creating global professionals. [Website]. Retrieved from <http://www.thelanguageflagship.org/>
- 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.
- Ullman, M. (2013). Declarative/procedural model. In Robinson, P. (Ed.) *Routledge Encyclopedia of Second Language Acquisition* (pp. 160-164). New York, NY: Routledge.

NOTES

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 second language acquisition research; but this is beyond the scope of the current exploratory study.
2. T-units are defined as sentences or independent clauses along with any accompanying dependent clauses.
3. 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 paper.
4. 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.