AN EXPLANATION OF THE EFFECTIVENESS OF WRITTEN CORRECTIVE
FEEDBACK IN SECOND-LANGUAGE ACQUISITION

by

Jason Wagner
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Linguistics

Committee:

_____ Director

_____

_____

_____ Department Chairperson

_____ Program Director

_____ Dean, College of Humanities
and Social Sciences

Date: _____ Spring Semester 2016
George Mason University
Fairfax, VA

An Explanation of the Effectiveness of Written Corrective Feedback in Second-Language Acquisition

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Jason Wagner
Master of Arts
George Mason University 2009
Bachelor of Arts
Campbell University, 1998

Director: Douglas Wulf, Professor
Department of Linguistics

Spring Semester 2016
George Mason University
Fairfax, VA

DEDICATION

This is dedicated to my loving wife Coco and our two wonderful boys Aiden and Dylan.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

vi

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Article .... art
Capitalization .... cap
Fragment .... frag
Kansas English Language Proficiency Assessment.............................. KELPA
Markedness Differential Hypothesis .... MDH
Masculine/Feminine Agreement .... mf
Preposition .... pr
Punctuation .... p
Run-on .... ro
Second Language... L2
Second Language Acquisition .... SLA
Singular/Plural .... num
Skill Acquisition Theory..SAT
Spelling ....sp
Subject-Verb Agreement .... s-v
Verb Tense .... vt
Word Form .... wf
Written Corrective Feedback ....WCF
Wrong Word .... ww

ABSTRACT

AN EXPLANATION OF THE EFFECTIVENESS OF WRITTEN CORRECTIVE
FEEDBACK IN SECOND-LANGUAGE ACQUISITION

Jason Wagner, Ph.D.

George Mason University, 2016

Dissertation Director: Dr. Douglas Wulf

The purpose of this study is to provide a theoretical explanation for the

effectiveness of Written Corrective Feedback (WCF) in increasing second-language (L2)

students' grammatical accuracy. WCF is examined via Skill Acquisition Theory (SAT)

in order to account for uneven patterns of its effectiveness. As the study demonstrates,

WCF is effective at increasing accuracy for some grammatical categories, but not for

others.   This SAT analysis also clarifies the seemingly contradictory findings in the

literature regarding WCF's overall effectiveness. Under SAT, for an instructional

strategy for grammar to be effective, it must: (1) impart adequate and clear declarative

knowledge of grammatical constructions and (2) provide a practice procedure whereby

L2 students can proceduralize and eventually automatize grammatical knowledge. WCF

that can accomplish these two goals is effective at increasing L2 students' grammatical

accuracy over time, whereas WCF that cannot is ineffective even when used over a substantial span of time.

The study has two major components. First, a quantitative study examines the rate of L2-English students' successful corrections on final drafts in response to WCF provided by the teacher in their first drafts (i.e., evidence of uptake). WCF in this study is a popularly used variety of indirect WCF. Descriptive statistics demonstrate that students corrected errors designated as *simple* (which involved a binary choice), but they either ignored feedback or made incorrect revisions to errors designated as *complex* (which involved a non-binary choice). Statistical analysis of the data demonstrates a statistically significant difference for WCF effectiveness on simple versus complex grammatical categories.

The second phase of the study examines WCF effectiveness for improving L2-Spanish students' grammatical accuracy (i.e., evidence of proceduralization) on new written works over time. The same type of indirect WCF was used as in the English phase of the study. The Spanish study was important in order to confirm the binary versus non-binary contrast by looking especially at idiosyncratic and non-idiosyncratic binary contrasts in addition to non-binary contrasts. Using a one-way repeated measures analysis of variance (repeated ANOVA), it was found that the grammatical accuracy improved for experimental group participants compared to control group participants for non-idiosyncratic binary contrasts, but not idiosyncratic binary contrasts. As with the previous study, WCF was also not successful for non-binary (i.e. > binary) contrasts. Thus, constructions designated as *simple* (i.e., treatable via this type of WCF) are only

those involving non-idiosyncratic binary contrasts. Constructions designated as *complex* (i.e., not treatable via this type of WCF) must be expanded to include not only non-binary contrasts but also grammatical paradigms with idiosyncratic binary contrasts.

This combination of quantitative and qualitative analyses of multiple sources of data from two different L2-language learning groups provides strong empirical evidence in support of SAT analyses of L2 instructional strategies like WCF. More precisely, these findings elucidate the reasons for WCF's variable effectiveness, providing a predictive utility for when L2 instructors can expect WCF to work or not. Specifically, only constructions with two possible choices with a non-idiosyncratic distribution improve over time through the use of WCF, such as the singular-plural contrast of subject-verb agreement in English. However, for a binary choice for which the two options have an idiosyncratic distribution, WCF is not effective for improving the whole grammatical category but can improve usage item-by-item, such as the Spanish masculine/feminine contrast for individual words. Lastly, constructions with more than two possible choices (i.e., non-binary) are also not responsive to WCF such as the three-way article choice in English of definite, indefinite, or null article. These findings demonstrate that WCF is effective at imparting sufficient declarative knowledge of whole grammatical paradigms only for binary constructions. Also, findings indicate that WCF could be effective at imparting knowledge of binary, idiosyncratic constructions, but only one item at a time.

These findings are important to the study of WCF because they provide a detailed, linguistic explanation of how WCF actually works instead of simply analyzing its overall

effectiveness and declaring it either to be effective or to be ineffective on the whole, which has so often been done in the literature on WCF. Under the SAT analysis used herein, *effectiveness* is defined and quantified according to its ability to impart declarative knowledge. This approach explains WCF's variable effectiveness, clarifies apparently contradictory empirical findings in the WCF research, and provides an objective measure of grammatical complexity that can be utilized in examinations of other types of corrective feedback and other L2-teaching strategies.

CHAPTER ONE: INTRODUCTION

Statement of the Problem

While second-language (L2) pedagogy and approaches to second-language

acquisition (SLA) have experienced an overall shift from form-focused (e.g. grammar

oriented) to meaning-focused (e.g., communicative) methodologies over the past several

decades (Sheen R. , 2003; Afitska, 2015), L2 learners are still under pressure to produce

grammatically accurate speech and writing in order to be successful in both academic and

professional pursuits (Ferris, 2006; Ellis R., Loewen, & Basturkmen, 2006; Afitska,

2015). One practical reason to avoid errors of form is to do well on language proficiency

assessments, which typically test both mastery of form and meaning.

However, since form errors can persist and fossilize even after years of L2 input

(White, 2003; Han, 2013), many L2 instructors have begun to reintroduce form-focused

strategies to supplement meaning-focused instruction with more explicit training in

grammar (Lightbown, 1992; Smith, 1993; Lightbown, 1998; Zephir, 2000; Loewen,

2005; Afitska, 2015). Zephir (2000) points out that "findings from classroom research

have begun to indicate that both form *and* meaning are important to the language

acquisition process, since neither one alone leads to complete second language

acquisition" (p. 20). The current study examines the effectiveness of Written Corrective

Feedback (WCF), a commonly used L2-teaching strategy that focuses L2 students'

1

attention on form in their own written work, from the perspective of theoretical linguistics. The analysis of WCF is conducted according to the principles of the theory of SLA known as Skill Acquisition Theory (SAT), (Anderson J. , 1976; McLaughlin, 1987; DeKeyser, 1997; 1998; 2001; 2007) a theory that postulates a strong connection or interaction between declarative and proceduralized language knowledge.

The use of WCF for teaching grammar in L2-writing classrooms is pervasive, despite controversies over its effectiveness in actually making L2 students more accurate writers (Truscott, 2007; Sheen Y., 2010; Shintani & Ellis R., 2013). Indirect WCF, a type of WCF commonly used in L2-writing classrooms, is an instructional procedure whereby errors in L2 students' writing are not directly corrected, but rather "flagged" with proofreading marks (e.g., *s-v* for a subject-verb agreement error, *vt* for a verb tense error, etc.). This feedback is accompanied by a correction symbol key that provides students with the name of the grammatical category the symbol represents, possibly a metalinguistic explanation of the structure, and an example of the error/symbol relationship. Students are then challenged to correct errors in their writing based on the feedback provided by the symbols and the explanations/examples from the correction symbol key in subsequent drafts of the same writing sample. It is hoped, and sometimes just assumed, that having L2 students revise their written work with the guidance of WCF correction symbols will result in decreasing rates of written errors over time (i.e., improved grammatical accuracy and increased grammatical competence). Henceforth, use of the term WCF is to be understood to refer to this commonly used variety of indirect WCF, as it is just this type of WCF investigated in this study.

Although only indirect WCF is examined in the study, direct WCF can be explained here briefly to note how it contrasts with indirect WCF. Direct WCF simply provides the correction, and students make the correction in subsequent drafts. There are no symbols to decipher with direct WCF. Some studies examine the relative effectiveness of indirect and direct WCF (Lalande, 1982; Robb, Ross, & Shortreed, 1986; Semke, 1984; Chandler, 2003; Ferris, 2006) while other examine the effectiveness of sub-types of each type. A brief review of the varied types is covered in subsequent chapters because results from the current study indicate that an SAT analysis and the measure of grammatical complexity used herein could prove instructive toward our understanding of the relative effectiveness of indirect and direct feedback types. Indirect WCF on binary errors is tantamount to providing the correct answer (e.g., "John eat" is marked with the symbol *s-v*, and there is only one option left for correct usage- "John eat<u>s</u>.") because "the other option" is the correct revision. This similarity could provide a research approach for examining the relative effectiveness of the two types of feedback.

This is akin to the ease of supplying amended answers to a previously graded true/false test where the location of missed questions has been indicated by a teacher. There is only one possible correction when the choice is a binary one. By contrast, supplying amended answers in a multiple-choice test involves more uncertainty since any of the remaining choices could be the correction the L2 student is attempting to provide. In this way, indirect WCF, by its very nature, can be seen to supply different levels of declarative knowledge feedback depending upon how many other correction options exist under a given grammatical construction's paradigm.

Research on WCF has been extensive, but perhaps in some cases too broad in scope and, in other cases, too narrow. Many studies just broadly ask the question *Is WCF effective, yes or no*? (Fathman & Whalley, 1990; Kepner, 1991; Ferris, 1999; Ashwell, 2000; Ferris & Roberts, 2001; Ferris, 2004). Such studies fail to reveal if WCF is indeed effective in addressing all kinds of error patterns and, for various errors, whether this effectiveness is uniform or uneven. Such broad examinations have provided only an indistinct picture of the operation of WCF for improving particular grammatical constructions. As I describe later, such evaluations of only the *overall* usefulness of WCF have plausibly led to apparently contradictory findings on WCF effectiveness.

On the other hand, some recent research on WCF has examined its use so narrowly, focusing on only one grammatical construction or on a subset of rules within a construction, that these pinpoint findings are neither applicable to the L2 classroom nor instructive toward our understanding of how WCF helps students to acquire or improve their use of complete grammatical systems (Sheen Y., 2007; Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2010; Ellis R., Sheen Y., Murakami, & Takashima, 2008; Sheen Y., 2010; Shintani & Ellis R., 2013).

The narrow and broad approaches to the study of WCF exemplify the differences between SLA researchers' and L2-writing researchers' differing research agendas and methodologies. SLA and L2-writing researchers begin their examinations of WCF by asking very different questions. While SLA researchers focus on WCF's capacity to expedite long-term acquisition of particular linguistic features, L2-writing researchers want to know if WCF can help their students be more accurate and better writers overall.

Of course, SLA researchers' narrower focus allows for tighter methodological controls compared to their L2-writing counterparts, but findings from L2-writing research is more relevant to the development of L2 pedagogy.

This study attempts to combine the strengths of the two approaches to the study of WCF by implementing strict methodological controls but also preserving the applicability of results to the L2 classroom. Arguably, the L2 classroom is the only place where WCF will be used to improve L2 learners' accuracy, so studies should be designed with this condition in mind. Moreover, this study utilizes an SAT analysis of WCF to provide answers for *how* and *why* WCF works as opposed to simply demonstrating that it works or does not work in particular situations.

Purpose of the Study

The current study applies SAT (Anderson J. , 1976; McLaughlin, 1987; DeKeyser, 1997; 1998; 2001; 2007.) to the study of WCF. SAT regards language acquisition as a specific instance of the more general phenomenon of skill acquisition. According to SAT, skill acquisition typically begins when knowledge about the skill (declarative knowledge) is imparted to the learner (e.g., "Hold the baseball bat all the way back, swing the bat straight across, and follow through."). Then, the learner must implement the declarative knowledge via extensive practice in order to build procedural knowledge of the skill (e.g., practicing the implementation of the imparted declarative

knowledge by actually swinging at baseballs in the batting cage extensively). As practice continues, the skill is developed, and the more one practices, the better the learner can become at the skill. Ultimately, procedural knowledge can become internalized such that it is automatic and unconscious (e.g., swinging the bat correctly and effectively without consciously thinking about the declarative knowledge). We see this type of skill learning occur with athletes, musicians, skilled laborers, and, indeed, L2 learners.

This is not to say that the acquisition of all language knowledge or skills derives via a skill acquisition paradigm. Language skills can develop as a result of many factors (e.g., L2 input from music and television, reading in the L2, conversations with native speakers of the L2, etc.). However, SAT claims and results from this study confirm that overt declarative training and practice can play a role in increasing proceduralized knowledge of forms.

An SAT analysis is used here for two reasons: (1) Wagner (2013) demonstrated that a link between conscious grammatical knowledge imparted via a teaching strategy (i.e., WCF) and unconscious grammatical production in new writing samples can exist, with a few stipulations, and (2) the principles of SAT can be seen in many other facets of SLA.

Indeed, the present study joins a growing body of literature that has appealed to SAT for an explanation of the development of specific grammatical structures in SLA. DeKeyser (1997) demonstrates that proceduralization of explicitly learned morphosyntactic rules can occur via practice of L2 grammatical rules and that successfully learning individual morphosyntactic rules is highly skill specific. Nagata

6

(1998) demonstrates that declarative knowledge of Japanese honorifics can be better developed via output practice (i.e., productive tasks such as speaking and writing) than input practice (i.e., passively receiving instruction on grammatical rules). Bird (2010), in a longitudinal study of the effects of distributed and massed practice schedules on the learning of L2-English syntax, discovers that distributed practice schedules are more effective in helping L2-English students proceduralize correct English syntax. Moreover, the benefits of distributed practice for L2 syntax were shown to be similar to those reported for other skills in the experimental psychology literature. Rogers (2011) demonstrates that increase practice time leads to automatization of verbal morphology. Kim, Lapointe, and Stierwalt (2012) find that varied combinations of feedback and practice lead to varying levels of effectiveness for the retention of novel phonetic productions of Korean phrases by native English speakers. Findings from these studies and many others support an SAT approach to examining SLA.

Observing WCF under the paradigm of SAT, it is possible to better understand how WCF may plausibly work and also how it sometimes fails to be useful. Simply stated, SAT maintains that WCF would be useful whenever it can contribute to the learner's declarative knowledge of grammatical constructions. The relevant declarative knowledge, in this case, is a conscious grasp of how the grammar works. Then, armed with this information, learners practice the forms by correcting them in subsequent drafts of the same writing sample. This practice acts as a bridge for a transition from the declarative knowledge provided by the WCF to a procedural knowledge of these same

forms. Procedural knowledge here refers to an ability to use the grammar accurately in actual language production, in this case academic writing.

This is how the concept of proceduralization is operationalized in the current study. Accuracy changes of experimental group participants (i.e., receive WCF) are compared to accuracy changes of control group participants (i.e., do not receive WCF) over time. While writing is certainly more of a conscious productive task than unrehearsed speech, changes in accuracy can be used to demonstrate that declarative knowledge has transitioned to proceduralized knowledge in the writing modality. Experimental participants who outperform control participants in accuracy increases over time demonstrate this transition. They begin to reduce errors significantly on certain constructions even though they are unaware that their accuracy changes are being examined and analyzed.

The studies herein are not designed to demonstrate complete automatization of knowledge, however. This would require a timed writing test of some kind, possibly on a computer, where production of various constructions could be observed in real time to determine the speed with which participants supply correct structures. For example, if participants would supply the correct subject-verb agreement markers (i.e., *s* or null) immediately to all present tense, third person verbs, we could claim that the knowledge was automatized. The current study, however, is designed to determine if a connection between declarative and proceduralized knowledge can be established. It cannot demonstrate complete automatization.

An SAT analysis allows us to define precisely what we mean by "effectiveness" of WCF. WCF is effective when it does what practitioners assume it should do, which is to impart sufficient declarative knowledge of a grammatical construction to permit useful practice in the use of the construction to begin, and it is not effective otherwise. SAT predicts that the effectiveness of WCF is fundamentally dependent upon what declarative knowledge is associated with a given grammatical construction, purely a matter of descriptive theoretical linguistics, and is not just a question of language teaching approach, student motivation, or any other consideration merely of pedagogical practice. The structure of the language itself is what determines when WCF works or does not work.

According to the preliminary findings of Wagner (2013) (seen in more detail below), indirect WCF is only effective at increasing accuracy for grammatical constructions that are linguistically *simple* (i.e., binary constructions). A simple construction is one for which the declarative knowledge of the grammatical pattern can straightforwardly be induced from the WCF alone. Initially, this definition of grammatical complexity/simplicity appears to suffer from the same circular reasoning that many pedagogical approaches to defining grammatical difficulty suffer from (i.e., The grammar is difficult because it is hard to learn, and it is hard to learn because it is difficult). However, the simple/complex distinction here is also based on an analysis of constructions' inherent linguistic complexities and empirical findings that have confirmed the binary vs. n-ary (n > 2) hypothesis for evaluating WCF effectiveness.

Findings corroborate that when WCF imparts decipherable declarative knowledge in the form of a repair strategy that can be clearly applied in subsequent writing samples, the correction process (i.e., the practice of making changes to flagged errors in a subsequent draft based on WCF feedback) can be expected to eventually move learners from declarative knowledge of a grammatical rule to a proceduralized knowledge of that rule (as measured by their increases in accuracy for particular grammatical forms over time).

When WCF by itself fails to impart the necessary declarative knowledge for correct grammatical usage for constructions that this study will describe as complex, making corrections to subsequent writing is simply inconsequential. L2 writers do not obtain declarative knowledge of the construction, cannot begin real practice toward the development of procedural knowledge over time, and thus will not demonstrate proceduralization of these constructions, even over an extended period of time of ongoing practice. In Wagner (2013), treatment group participants demonstrate transitions from declarative to proceduralized knowledge for the simple construction of English subject-verb agreement and English singular/plural. These constructions involve a binary choice (add "s" or do not add "s"). When L2-English students encounter this flagged error, they can deduce that "the other" option is the correct option. Crucially, there are no other options for correct usage, and the rule applies relatively uniformly in the location of a third person singular verb and a plural noun.

They do not demonstrate such transitions for the complex construction of English verb tense or article usage. Correct verb tense usage depends on knowledge of twelve

verb tenses, the correct use of a large amount of linguistic materials, and knowledge of the underlying semantics of the twelve tenses. Articles involve only three options (definite, indefinite, and null articles), which already exceed the binary limit for WCF effectiveness. However, these three options do not correspond to a one-to-one form and meaning relationship, which complicates usage even further. L2-English students who encounter these types of flagged errors receive no information on correct usage, only that they have made an error. Attempts to repair the error are generally only guesses, and these guesses cannot confirm declarative knowledge of a construction, nor can this type of practice move L2-English students' knowledge of constructions from declarative to proceduralized knowledge.

Findings from Wagner (2013) prompted the following questions: (1) Is the simple/complex contrast simply a matter of a binary/non-binary distinction, and (2) is WCF effective at an even more limited range based on an idiosyncratic/non-idiosyncratic contrast? The study herein helps to answer these questions and provides a strong empirical evidence for a more discerning use of WCF.

As the findings of this study will demonstrate, an assessment of WCF based on its capacity to impart useful declarative knowledge appears to be an effective approach to assessing WCF effectiveness. It allows us to designate for which points of grammar WCF is effective and for which it is not. Furthermore, the cutoff in effectiveness appears to be determined entirely on linguistic considerations. Specifically, WCF is effective for grammatical forms whose rules can be sufficiently explicated for the learner via WCF so as to permit acquisition of declarative knowledge of the form as delineated under SAT.

This, in turn, permits the learner to practice use of the form and eventually gain the necessary procedural knowledge under SAT. In short, as the study results will reveal, WCF is shown only to be effective with helping learners acquire grammatical patterns involving a binary choice that is non-idiosyncratic in its distribution. Indeed, for such points of grammar, WCF is extremely helpful. By contrast, WCF is not effective whatsoever for non-binary grammatical paradigms and has only limited success with idiosyncratic, binary forms. That is, WCF is shown not to be effective for helping learners acquire grammatical patterns involving non-binary choices, and it can only impart declarative knowledge of idiosyncratic, binary constructions item-by-item.

In addition, findings from the current research study support the applicability of SAT in SLA. Findings support that a strong connection/interaction can exist between declarative and proceduralized knowledge, supporting the strong-interface position in SLA (Sharwood-Smith, 1981; Bialystok & Ryan, 1985; Anderson J. , 1983; 1990; 2005; Dienes & Perner, 1999; White & Ranta, 2002; Ellis R. , 2009a). The findings are thus evidence against the non-interface position (Krashen, 1981; Krashen, 1982; Krashen & Terrell, 1983; Schwartz, 1993; Paradis, 1994; 2009; Hulstijn, 2002; Ellis R., 2005; 2009b; Ellis N. , 2005), which claims that the two types of knowledge do not interact at all in SLA. Indeed, SAT appears to have a great deal of explanatory power regarding how SLA might be postively impacted in instructed environments.

Research Questions and Hypothesis

Based on Wagner's (2013) initial examinations of WCF effectiveness on increasing L2-English writing students' grammatical accuracy, this study advances the following research questions:

> **RQ 1:** Does WCF and the correction process aid in the transition from declarative to proceduralized knowledge of specific grammatical constructions in the writing modality? And if so, for which constructions is this transition observed?

> **RQ 2**: Given the answers to RQ1, is there a grammatical complexity threshold beyond which WCF will no longer be operative in supplying the declarative knowledge necessary for L2-writing students to practice and eventually proceduralize?

This approach to the study of WCF provides a more detailed examination of its effectiveness/ineffectiveness and allows us to quantify precisely WCF effectiveness according to an objective and falsifiable measure of grammatical complexity seen in the research hypothesis below:

> **Hypothesis**. WCF is only effective in imparting declarative knowledge thereby increasing linguistic accuracy (proceduralization) for grammatical constructions that include a binary choice of options for correct usage (binary vs. n-ary ($n > 2$)).

The research hypothesis provides delineation between simple and complex forms according to an objective measure of complexity based on the SLA theory of SAT. It allows us to examine the intricacies of WCF success and failure as opposed to broadly claiming that it is either effective or ineffective. Moreover, results supporting the hypothesis reinforce SAT as an informative way to depict the impact of instruction on SLA. Additionally, it was important to determine if the simple constructions for which WCF is effective would be the binary-choice constructions and if the complex constructions are the n-ary-choice (non-binary) constructions. As the results of the study

reveal, in addition to the n-ary-choice constructions, binary constructions with an idiosyncratic distribution were additionally found to be complex, in that these grammatical categories as a whole were not amenable to improved acquisition via WCF.

Significance of the Study

The results of the current study largely support the research hypothesis, demonstrating a grammatical complexity threshold beyond which WCF will no longer be effective in helping students to acquire complete grammatical systems. Results also augment the applicability of SAT in SLA and the strong-interface position for SLA, demonstrating that a strong interaction *can* exist between declarative and proceduralized grammatical knowledge in L2 acquisition. Also, examining the effectiveness of WCF for both L2-English and L2-Spanish learners broadens the applicability of the findings to more than just one language-learning group. Indeed, this WCF study is the only one of its kind, examining WCF effectiveness cross-linguistically in order to compare WCF effectiveness on similar and dissimilar grammatical constructions between the two languages. Spanish was not simply selected because it is a widely taught second language in the United States. Crucially, Spanish provides us with an example of a binary error pattern that exhibits idiosyncrasy, allowing us to see if the research hypothesis provided above is stated strongly enough.

Before discussing the study, we turn now to examine the most recent issues in WCF scholarship. In the following chapter, we survey both empirical studies of WCF and the theoretical paradigms applied to the current study.

CHAPTER TWO: LITERATURE REVIEW

Grammatical Complexity/Difficulty

An SAT analysis of WCF hinges on defining grammatical complexity accurately and in such a way that varying levels of complexity can be identified and compared. Indeed, it is varying levels of grammatical complexity that determines WCF effectiveness and the ability of L2 students to successfully acquire or not acquire grammatical systems via the use of WCF. The lack of grammatical complexity definitions is precisely what has lead to contradictory findings in the literature on WCF, which will be seen in more detail below.

Defining grammatical difficulty or establishing an objective measure of what constitutes a simple or complex grammatical construction in L2 learning has proven to be an extremely challenging undertaking. Vast differences in grammatical complexities do not lend themselves to simple language internal or cross-linguistic comparisons of difficulty. Even superficially similar grammatical patterns differ significantly in complexity within languages and cross-linguistically. Moreover, approaches to defining grammatical difficulty vary widely and are based on a number of divergent and often contradictory criteria.

For example, while one researcher claims English subject-verb agreement is simple based on the number of alternative forms (add –s or do not add –s to the base form

of a verb) (Krashen, 1982), another claims it is complex based on the construction's processing demands (the long distance between the subject and the end of the verb) (Ellis R., 1990). Another concludes that the construction is complex but bases the assessment on yet another criterion, its highly syncretic nature (the combination of three abstract concepts— present tense, third person, singular) (DeKeyser, 1998). Not only is there a lack of consensus on what constitutes simple and complex, but there are also vast inconsistencies in how criteria have been applied in empirical studies attempting to distinguish simple constructions from complex constructions.

However, the principle problem with grammatical difficulty assessments is that the criteria used in quantifying complexity are often too narrow and inadvertently discount significant factors in the evaluations of what makes a structure simple or complex. Researchers have been prone to applying one difficulty criterion (e.g., number of transformations from base to surface form) to a large variety of grammatical constructions and then categorizing these structures as simple or complex according to that single criterion. Consequently, structures that are notoriously difficult for L2 English learners to acquire (e.g., English articles) are categorized as "simple" because, according to that one criterion (i.e., require no transformations), they are not complex. Approaches to defining grammatical difficulty generally fall into one of three categories: psycholinguistic, linguistic, and pedagogical (Spada & Tomita, 2010), as seen in detail below.

Psycholinguistic definitions. The psycholinguistic approaches focus on developmental readiness and predictable stages of acquisition as the determinants of

grammatical difficulty. Many empirical studies have demonstrated support for acquisition of L2 grammar in English based on developmental readiness including findings from studies of English morphological and syntactic features (Meisel, Clahsen, & Pienemann, 1981; Pienemann, 1989; Pienemann, Johnson, & Brindley, 1988), studies of relative clauses (Keenan & Comrie, 1977; Doughty, 1991; Izumi & Izumi, 2004; Ammar & Lightbown, 2005), studies of tense and aspect (Anderson & Shirai, 1996; Bardovi-Harlig, 2000; Shirai, 2004), studies of negation (Dulay, Burt, & Krashen, 1982), and studies of English question forms (Mackey, 1999; Mackey & Philip, 2011; Spada & Lightbown, 1993). However, the developmental readiness or staged acquisition approaches suffer from circular reasoning: The construction is difficult because it is learned later, and it is learned later because it is difficult.

Linguistic definitions. Linguistic definitions of complexity focus on things like the inherent complexities of structures, L1 transfer, communicative force, and saliency and frequency. It is generally assumed that the more complex the rules of a grammatical construction, the more difficulty an L2 learner will have in learning or acquiring the construction. A number of researchers have attempted to define and assess measures of grammatical difficulty by examining the inherent complexity of structures (Hulstijn & de Graaff, 1994; Givon, 1991;1995; Robinson, 1996; DeKeyser, 1998; Doughty & Williams, 1998; Housen, Pierrard, & Van Daele, 2005; Spada & Tomita, 2010). Hulstijn & de Graaff (1994) define grammatical complexity by the number of transformations a construction undergoes for correct usage while DeKeyser (1998) tabulates the number of forms, meanings, and form-meaning relationships to determine levels of complexity.

L1 transfer can also be used to define grammatical complexity (Lado, 1957; Stockwell, Bowen, & Martin, 1965; Eckman, 1977; White L. , 1991; Master, 1991; Gass & Selinker, 1992; Yip, 1995; Odlin, 1989; 2003; Spada & Lightbown, 1999; Lightbown & Spada, 2000; Hinkel, 2002; Izquierdo & Collins, 2008). Here grammatical complexity is determined by differences between the learners L1 and L2, and L1 transfer can either facilitate (positive transfer) the acquisition of grammatical structures or inhibit it (negative transfer).

For example, an L2-English student who speaks Spanish as her L1 may find the concept of verb tense and the production of temporal aspect via verb morphology easier to grasp than an L2-English student who speaks Chinese as her L1. Grammatical complexity/difficulty is relative under an L1 transfer paradigm and depends on the L2 learners' L1. Spanish requires the use of morphology for tense and aspect in much the same way English does. This similarity can lead to positive transfer, and L1-Spanish speakers may encounter fewer difficulties with acquiring L2-English verb tenses. However, L1-Chinese speakers generally establish time differently, usually with the use of adverbs at the beginning of sentences (e.g., *gloss*: Yesterday, I go. Today, I go. Tomorrow, I go.) This difference can make learning L2-English verb tenses relatively more difficult for L1-Chinese speakers who have less experience with verbal morphology and the subtle semantic differences English verb tenses represent.

Eckman (1977) goes beyond the idea of relative difficulty or complexity between L1s and L2s by introducing the Markedness Differential Hypothesis (MDH). The MDH does not only rely on a comparison of L1 and L2 to predict difficulty in L2 learning.

"This notion of difficulty must be independent of any give language (i.e., must be universal), and must be valid on grounds which are independent of the facts surrounding SLA" (p. 320). Eckman defines difficulty according to typological markedness and implication relations. The MDH relies on comparisons of the L1, the L2, and markedness relations in universal grammar. These markedness relations are defined as follows:

> A phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A.

Defining complexity independently and universally is more objective. According to the MDH, structures in the L2 that are different and more marked than the L1 will be difficult to learn. Structures that are different and less marked than the L1 will be easy to learn.

The communicative value of a grammatical structure has also been examined as a criterion to determine grammatical difficulty (VanPatten, 1996; 2002; 2004; 2007; Harrington, 2004; Spada & Lightbown, 2008). Where a structure or an error is important to meaning, it is said to be easier to learn or to correct because learners focus on meaning, not grammar, while processing input. Findings confirming communicative value effects on levels of difficulty have been used to support meaning-focused, communicative language teaching strategies. It is a paradigm firmly rooted in the non-interface hypothesis of SLA (seen in more detail below), which claims that explicit and implicit knowledge of grammatical structures do not interact. Under such a paradigm, inherent complexities, markedness relations, saliency, and frequency do not matter. Structures

that are important to meaning will be learned more easily and possibly sooner than structures that are not important to meaning. Consequently, proponents of this approach to defining grammatical complexity claim explicit instruction on form is not necessary and is possibly useless. Communicative language strategies such as having L2 students discuss events in the simple past tense and providing recasts on errors facilitates acquisition of proper simple past usage better than explicitly teaching simple past rules.

Linguistic definitions of difficulty also focus on saliency (Skehan, 1998; Doughty & Williams, 1998; Goldschneider & DeKeyser, 2005; Collins, Trofimovich, White, Cardoso, & Horst, 2009; DeKeyser, Alfi-Shabtay, & Ravid, 2010) and frequency (Ellis N. , 2002; Ellis R., 2009b; Gass & Mackey, 2002; Goldschneider & DeKeyser, 2005) as criteria for assessing grammatical structures' complexity. Generally speaking, more salient forms are more likely to be noticed in the input because they stand out or are easier to hear. Consequently, the claim is that the more salient a form is, the easier that form is to learn. Dulay & Burt (1978) claim that "perceptional saliency is an input factor that has not yet been precisely defined" (p. 73). Although saliency is recognized as an important factor in SLA, it can be difficult to operationalize precisely what saliency actually is. Goldschneider & DeKeyser (2005) attempt to break down saliency into three sub-factors: the number of phones in the functor (phonetic substance), the presence/absence of a vowel in the surface form (syllabicity), and the total relative sonority of the functor. Still, attempting to define categorically which forms are more salient than others can be a difficult proposition.

More frequent forms are also more likely to be noticed in the input, thus making these constructions easier to learn, though Master (1994) argues convincingly that the ubiquitous nature of English articles (i.e., some of the most frequently used structures in English) actually makes them harder to learn. Of course, one could also argue that articles are not very salient in the input because they are generally produced in unstressed syllables. It can be difficult to tease apart saliency and frequency effects, especially if a form is both frequent and salient (e.g., progressive *ing*).

Pedagogical definitions. Pedagogical definitions of complexity are generally explained in terms of students' performance in producing grammatically accurate speech or writing. This approach does not depend on a comparison of students' performance and a set scale of their developmental readiness or predictable stages of acquisition like the psycholinguistic definitions above. Under this approach, difficult constructions are identified by teachers based on their observations of L2 students' performance and systematic errors (Robinson, 1996).

For example, an L2 teacher observes that students generally perform well in their use of simple past tense, but these same students perform badly for an extended period of time in their use of the present perfect tense. Under a pedagogical approach to defining grammatical complexity, the present perfect tense is designated as more complex/difficult than the simple past tense. Other researchers define grammatical difficulty from the learners' perspective (Scheffler, 2009) asking students to rate constructions as simple or complex and comparing their performance against said students' own assessments of grammatical levels of difficulty.

This Study's Definition of Grammatical Complexity

The current research study defines and assesses grammatical complexity according to very specific criteria that relate directly to an L2-teaching strategy, specifically WCF. It combines linguistic and pedagogical approaches to defining grammatical complexity. Primarily, grammatical constructions that have only two, non-idiosyncratic options for correct usage are defined as simple while constructions that have more than two options for correct usage or a binary, idiosyncratic option are defined as complex. This approach to defining complexity relies on the inherent complexity of structures. It is applied to similar and dissimilar constructions in both English and Spanish, as WCF effectiveness is assessed for both L2-English and L2-Spanish learners. Whereas English and Spanish singular/plural require only the addition or deletion of "s" morphology in order to correct an error, verb tense in both languages is far more complex. Furthermore, this cross-linguistic analysis allows for a comparison of English subject-verb agreement, designated "simple" by the research hypothesis, and Spanish subject-verb agreement, designated "complex" by the research hypothesis.

English subject-verb agreement depends on a simple binary choice for correct usage. In English, we add an "s" to present tense verbs when they are accompanied by a third-person, singular subject. Otherwise, the verb requires no "s" morphology. Spanish, on the other hand, is a morphologically rich language. Every subject type requires a different verb ending, as seen in more detail below, and declarative knowledge of these complexities cannot be imparted by WCF.

Because simple constructions have only two options to learn for correct usage, this allows declarative knowledge of these constructions to be induced from WCF alone. With these types of "simple" constructions, a WCF symbol not only points out and focuses students' attention on an error that has been made but also provides the correct answer (i.e., the "other" option). *Crucially, there are no other options within the construction for correct usage, even within varied semantic or pragmatic situations.* Indeed, the binary composition of these simple grammatical structures (e.g., add s or do not add s) coincides with the binary nature of WCF (i.e., correct or incorrect). Where there is no WCF, students assume there are no errors (i.e., correct). Where they receive WCF, an error has been made (i.e., incorrect). For simple constructions, providing the feedback is tantamount to providing the correct answer. The "other" option is correct. Clear declarative knowledge is imparted, and it can be practiced meaningfully.

For example, when L2 students receives a *num* symbol above a noun error, they either have either to add an "s" or delete an "s" to correct the error. If their original noun has no "s," the L2 students know to produce a noun with plural "s" morphology in the subsequent draft of the same writing sample. The student knows that the correction they have made is correct. There are no other options for correct usage. The certainty of this knowledge is fundamental to the successful implementation of WCF. It is precisely this certainty that allows for meaningful practice to begin, and this type of practice can lead students to proceduralization of the clearly imparted declarative knowledge. It is also this certainty that makes indirect WCF on binary forms similar to direct WCF, which simply

provides the correct answers in rough drafts for L2-writing students to copy in final drafts.

On the other hand, complex constructions are those that have more than two options for correct usage. When an L2 teacher provides WCF for a construction with more than two options for correct usage (e.g., *vt*- "verb tense"), the feedback points out the error without providing any decipherable declarative knowledge for what the correction should be or why the correct revision is correct. In fact, there is no way for the L2 student to know if he or she has made the right correction in the subsequent draft.

For example, when an L2 student encounters a *vt* symbol above a verb tense error, there is no decipherable declarative knowledge to learn from the symbol. Options abound for correct verb usage, and even when an L2 student chooses the correct revision, that student cannot know that the revision is correct. The uncertainty of what the correct revision is and the uncertainty of whether the correct revision was made renders WCF ineffective. Declarative knowledge is not imparted, and practice cannot bridge the gap between declarative and proceduralized knowledge.

Linguistic Perspectives on Declarative and Proceduralized Knowledge

The current research study assesses WCF effectiveness on how well it imparts declarative knowledge and how successfully the correction process aids in the transition from declarative to proceduralized knowledge. Theories of SLA differ on whether a connection between the two types of knowledge even exists. Consequently, a brief

discussion of SLA views on the two types of knowledge helps to put results from the current research study into perspective.

Theoretical perspectives on how declarative and proceduralized knowledge of languages are learned and how these two types of knowledge interact in L2-language learning vary considerably. These perspectives are germane to a discussion of WCF and SAT because SAT maintains that language learning can be characterized as a movement from the one type of knowledge (explicit) to the other (implicit). The two types of knowledge have been labeled in a variety of ways: *unanalyzed* versus *analyzed* knowledge (Bialystok & Ryan, 1985); *acquired* versus *learned* knowledge (Krashen, 1981); and *procedural* versus *declarative* knowledge (Anderson J. , 1983; DeKeyser, 1998). Despite the varied descriptive labels associated with the two types of knowledge (and the subtle theoretical positions they represent), declarative knowledge is generally regarded as explicit knowledge that is available during controlled processing in tasks like explaining a grammatical rule or in writing and editing. According to SAT, *effective* WCF must impart this type of knowledge.

Proceduralized knowledge, on the other hand, is implicit or automatized knowledge and is available for automatic, unconscious processing in spontaneous, unrehearsed speech or writing. Under an SAT analysis of WCF, increased grammatical accuracy in using a particular grammatical form implies proceduralization of that form. Both types of knowledge are relevant to an examination of WCF under SAT because, under such an analysis, WCF effectiveness is assessed through its usefulness in imparting declarative knowledge (explicit) that can be practiced and eventually transition into

25

proceduralized knowledge (implicit). Indeed, because the correction process (i.e. the practice of correcting the error in subsequent drafts of the same writing sample) is the same for all the grammatical categories, WCF effectiveness must be measured by its ability to communicate declarative knowledge of forms that can then be practiced fruitfully in order to build procedural knowledge. In the current study, the concept of proceduralization (i.e., the development of implicit knowledge) is operationalized by measuring accuracy increases over time on new writing samples. According to findings here, these increases only occur for experimental groups (receive WCF), not control groups (do not receive WCF), and they only occur for grammatical constructions that are designated as simple, not complex. As mentioned in Chapter 1, experimental and control group participants' performance in accuracy changes is how evidence of proceduralization is being measured.   Where experimental participants demonstrate accuracy increases relative to their control group counterparts, it demonstrates that WCF and the practice procedure of correction are instigating a transition from declarative knowledge to a proceduralized use of rules.

Some L2-language researchers argue that declarative and proceduralized knowledge do not interact at all as L2 learners acquire their L2, a position referred to as the non-interface position (Krashen, 1981; Krashen, 1982; Krashen & Terrell, 1983; Schwartz, 1993; Paradis, 1994; 2009; Hulstijn, 2002; Ellis R., 2005; 2009a; Ellis N. , 2005). However, this theoretical perspective on SLA cannot explain L2 students' abilities to proceduralize knowledge at least for some grammatical forms when WCF is utilized. That is, if no interaction is taking place between the two types of knowledge, it

becomes difficult to explain how L2-writing students significantly improve on the use of some forms but not others, as we see in the results of the current research study.

This is not to say that procedural competence cannot develop without overt declarative training. Certainly, proceduralization of linguistic knowledge can progress as a result of a wide array of input sources (e.g., reading in the L2, conversations with native speakers of the L2, listening to radio and television in the L2, etc.). However, research on L2 instructed environments and findings from this study confirm that explicit instruction on grammatical forms can positively impact accuracy and successful acquisition of grammatical patterns (Lightbown, 1992; Smith, 1993; Lightbown, 1998; Zephir, 2000; Loewen, 2005; Afitska, 2015). Generally speaking, L2 students who receive overt grammatical training outperform those who learn their L2 devoid of such training. The current study does not endeavor to examine all of the possible influences that could be at play in an L2-students' acquisition of linguistic constructions and attempts to control for such outside influences in order to examine the effects of WCF alone (see the sections labeled Study Design below). The focus of the current study is to demonstrate that the transmission of clear declarative knowledge via WCF can act as a foothold for meaningful practice of grammatical items to occur and that WCF effectiveness can be measured exclusively on its capacity to impart clear declarative knowledge. Once clear declarative knowledge is imparted, it can be practiced, and practice can bridge the transition to proceduralization of declarative knowledge.

Indeed, WCF provides some declarative knowledge feedback on grammar. If the non-interface position was correct in all instances of L2-language acquisition, WCF

should be uniformly ineffective. However, WCF has been shown to be effective in a variety of studies (Ferris, 1999; Ashwell, 2000; Ferris & Roberts, 2001; Fathman & Whalley, 1990; Ferris, 2004; Kepner, 1991; Wagner, 2013).

Conversely, proponents of SAT maintain a strong interaction between declarative and proceduralized knowledge, a position referred to as the strong-interface position (Bialystok & Ryan, 1985; Sharwood-Smith, 1981; Anderson J. , 1983; 1990; 2005; Dienes & Perner, 1999; White & Ranta, 2002; Ellis R., 2009b). Indeed this strong-interface position is generally grounded in SAT. Again, SAT takes a dichotomous view of declarative and proceduralized knowledge and views L2-language learning as a transition from declarative knowledge of particular grammatical constructions in a language to their procedural mastery (a process realized via practice). The strong-interface position better explains how the acquisition of grammatical constructions takes place *with the use of WCF* by providing a stipulation that all conditions are met for a transition of knowledge when using WCF (i.e. declarative knowledge is imparted, practice can occur, proceduralization follows). According to SAT, where these conditions are met, WCF will be successful. Where they are not met, WCF will fail.

In studies of WCF, the "practice condition" mentioned above is held constant because the practice is *precisely* the same for all grammatical categories under analysis (subject-verb agreement, singular/plural, verb tense, articles, prepositions, and masculine/feminine). For all constructions, practice equates to students attempting to decipher the meanings of the symbols for flagged errors and subsequently making corrections to these errors between their rough and final drafts. Proponents of WCF

assume that this practice will lead to increased grammatical competency for *all* the flagged errors. Because the practice of correction is a constant across all the categories, the current study need only focus on WCF symbols' ability to impart clear declarative knowledge. Where clear declarative knowledge is conveyed by a symbol, meaningful practice (i.e., students correcting flagged errors between rough and final drafts) occurs and proceduralization follows. Proceduralization is confirmed by participants' abilities to maintain knowledge of a form (as demonstrated by increased accuracy on forms) over extended periods of time. This study examines and measures proceduralization by comparing experimental participants' accuracy increases to control groups' accuracy increases over time. When experimental groups outperform control groups, it demonstrates that proceduralization has occurred as a result of WCF.

Finally, some researchers see only a weak interaction between explicit and implicit knowledge (Ellis R., 1994; 2005; Schmidt, 2001; Doughty & Williams, 1998; Norris & Ortega, 2000) where explicit knowledge helps L2 students to "notice the gap" (Schmidt, 2001) between their production and correct, native-like production. It could be argued that WCF helps students to "notice the gap," but this would only be possible where clear declarative knowledge is imparted for the purposes of distinct comparisons, and here again, WCF effectiveness would have to be measured based on its abilities to impart clear declarative knowledge. Where ambiguous declarative knowledge or no declarative knowledge is imparted, no distinct comparisons can be made, and a "gap" cannot be noticed.

Regardless, as long it can be demonstrated that students make progress in increasing their grammatical accuracy on some forms with the use of WCF, it will not matter if the strong-interface position or weak-interface position is correct. The current research study does not attempt to quantify the *amount* of interaction (weak or strong) between explicit and implicit knowledge. Rather, this study assesses the existence or nonexistence of a connection/transition between these two types of knowledge via practice.

Indeed, it would be extremely difficult for a study of this kind to measure the degree or amount of connection existing between explicit and implicit knowledge in an L2 learner. A scale of connection strength would have to be developed and examined to determine how much connection exists between participants' declarative knowledge of a structure and their procedural knowledge of that same structure. Measuring a degree of connection would require an initial examination and accounting of L2 students' levels of declarative knowledge (i.e., no declarative knowledge to full declarative knowledge). Such an examination would also necessitate the development of a scale of proceduralization (i.e., not proceduralized to fully proceduralized), and such a scale would be difficult to develop and test empirically. Only then could L2 students' level of declarative knowledge and level of proceduralized knowledge be compared to examine the strength of connection between the two. However, even when L2 students could demonstrate high levels of both types of knowledge on a particular form, it would still be difficult to verify that a high degree of connection was at play. Proponents of the non-

interface position could simply argue that high levels of both types of knowledge can and do exist separately from one another.

The objective of the current research study is to determine if there is an interface or not. Establishing that a connection can or cannot exist between the two types of knowledge must precede discussions of their degree of connection. Moreover, the complimentary design and findings of the two empirical studies herein examine the SAT transition process from beginning to end, and this approach helps us to observe the connection or lack of connection between the two types of knowledge. Specifically, the L2-English WCF Uptake Study examines the beginning of the transition process by demonstrating which symbols impart sufficient declarative knowledge for meaningful practice to begin. Where students make correct revisions to grammatical errors in subsequent drafts (i.e., positive uptake), we can assume that the WCF symbols impart sufficient declarative knowledge and that practice is consequential. Where students are unable to make correct revisions or make new errors to grammatical errors in subsequent drafts, (i.e., no uptake), we can assume that WCF symbols cannot impart sufficient declarative knowledge and that practice is inconsequential. This L2-English WCF Uptake Study examines the beginning of the transition process between the imparting of declarative knowledge via WCF and the practice procedure.

The L2-Spanish WCF study, on the other hand, examines the full transition process with a focus on the practice procedure and the examination of accuracy changes over time (i.e., proceduralization). Accuracy increases/error reductions is how we operationalize the concept of proceduralization for this study. Where experimental

participants demonstrate decreases in error rates (i.e., accuracy increases) compared to their control counterparts, we can assume that the practice procedure aided in transitioning knowledge from declarative to proceduralized knowledge of the forms under examination. Findings from the two studies reinforce one another regarding which grammatical constructions are and are not amenable to positive change via the use of WCF. Both studies demonstrate that it is the level of inherent complexity of constructions that must be accounted for when discussing potential connections existing between declarative and proceduralized knowledge (at least with the WCF examined in this study).

Broad Studies of WCF Overall Effectiveness

As noted previously, many WCF studies have simply examined whether WCF as a whole is effective, confirming that WCF is either effective or not at increasing *overall* writing accuracy (Fathman & Whalley, 1990; Kepner, 1991; Ferris, 1999; Ashwell, 2000; Ferris & Roberts, 2001; Ferris , 2004). Likewise, my initial study of WCF, Wagner (2013), found WCF to be effective overall when comparing experimental and control groups' abilities to reduce all error types over time (Corresponding ANOVA P Value = .004 with a statistical significance level of < .05) Still, some researchers have argued against the effectiveness of WCF on both theoretical and empirical grounds (Truscott, 1996; 1999; 2004; 2007; Polio, Fleck, & Leder, 1998), claiming that WCF is not only ineffective but even potentially harmful to L2 learners.

Arguments against WCF are based on a nativist view of L2-language instruction and acquisition. This view of SLA sets the foundation for the meaning-based approaches to L2 instruction and exemplifies the principles of the non-interface position of SLA. That is, explicit and implicit knowledge are separate, and explicit instruction on form will not positively impact L2 students' accuracy in unrehearsed speech and/or writing. Opponents of WCF argue that students are incapable of deciphering the meanings of WCF symbols or understanding and applying complex grammatical rules. As will be seen, this is a claim partially supported by findings from the current study. Opponents of WCF argue that a writing curriculum should focus on the higher order aspects of writing and that grammatical accuracy and acquisition will follow without explicit instruction on form. This is an argument firmly grounded in the non-interface position of SLA (Krashen, 1981; Krashen, 1982; Krashen & Terrell, 1983; Schwartz, 1993; Paradis, 1994; 2009; Hulstijn, 2002; Ellis R., 2005; 2009a; Ellis N., 2005), which claims that there is no connection between explicit and implicit knowledge of rules. A detailed review of the WCF debate is presented below.

Truscott (1996) initiated the debate over WCF's effectiveness. He argued:

… grammar correction has no place in writing courses and should be abandoned. The reasons are: (a) Research evidence shows that grammar correction is ineffective; (b) this lack of effectiveness is exactly what should be expected, given the nature of the correction process and the nature of language learning; (c) grammar correction has significant harmful effects; and (d) the various arguments offered for continuing it all lack merit (p. 328).

Truscott's (1996; 1999; 2004; 2007) main thesis has been that the use of WCF specifically and explicit grammar instruction generally do not work in making students better writers (i.e., not better in terms of either content or form). Also, Truscott makes the very strong claim that, of the varying types of WCF that can be employed in writing classes, none are effective and all should be abandoned: "There is no reason to think any of the variations should be used in writing classes, and there is considerable reason to think they are all misguided" (p. 329). Truscott has challenged WCF researchers to either find support for or evidence against his strong and experimentally falsifiable claims. However, his positions neither account for varying levels of complexity among different grammatical constructions nor WCF's variable effectiveness at imparting knowledge of said constructions. Truscott's call to do away with WCF altogether before accounting for this stipulation has been premature.

Truscott's review of methodologically unsound research on WCF has been a major contribution to the field. He has highlighted faulty methodology and conjecture reported as empirical evidence (Truscott, 2004), calling WCF studies that compare control groups and experimental groups and for statistical analysis of data that incorporates analyses of effect size, confidence intervals, and clear research hypotheses. He also calls for studies of WCF to use valid measurement instruments (e.g., authentic writing samples, not grammar exercises).

Ferris (2004) confirms Truscott's concerns regarding methodology in WCF research, demonstrating that of the more than thirty studies being reviewed and reanalyzed at the time, only six examined a correction group (experimental or treatment)

against a no correction group (control) (Semke, 1984; Fathman & Whalley, 1990; Kepner, 1991; Polio, Fleck, & Leder, 1998; Ashwell, 2000; Ferris & Roberts, 2001). Of these six, only two examined the effects of WCF over a significant period of time (e.g., weeks or months) (Kepner, 1991; Polio, Fleck, & Leder, 1998). Many of the previous studies of WCF effectiveness simply examined whether WCF improved final drafts relative to first drafts of the same writing sample.

Of the six studies that had experimental and control groups, four provide support for the effectiveness of WCF (Fathman & Whalley, 1990; Kepner, 1991; Ashwell, 2000; Ferris & Roberts, 2001), one is missing information and is inconclusive (Semke, 1984), and one provides support for Truscott's position, reporting no advantage for the group that received error correction (Polio, Fleck, & Leder, 1998). Consequently, as of 2004 we only had two studies (Kepner, 1991; Polio, Fleck, & Leder, 1998) that examine WCF effectiveness for the acquisition of grammatical rules, one demonstrating support for the use of WCF (Kepner, 1991) and one calling its usefulness into question (Polio, Fleck, & Leder, 1998).

In addition, Truscott's (1996) reanalysis and interpretation of Kepner (1991) contradicts Ferris' interpretation. Truscott claims that the results of Kepner's study do not show support for WCF effectiveness because of methodological flaws. Truscott recounts that Kepner did not have a true control group because one group received comments in their native language on sentence level errors while the other received feedback on content in the target language. Also, students were not required to do

anything with the correction feedback in the Kepner study, so it would be difficult to attribute findings to the use of WCF.

Truscott's critiques of WCF research methodology, while well-founded, should have led to a conclusion that findings on WCF effectiveness were inconclusive, not that the implementation of WCF should be altogether abandoned. The current study remedies the methodological flaws of past WCF research. They compare control and experimental groups over a significant period of time, analyze data for statistical significance, and use authentic writing samples from L2-English language learners and L2-Spanish language learners. The study includes clear research questions, a falsifiable hypothesis, and statistical analysis of the data collected. Most important, the current research represents a new approach to the study of WCF that accounts for grammatical complexity and explains findings within a theoretical framework of SLA.

Ferris's (1999) response to Truscott's (1996) various critiques of WCF and WCF studies touches on a research agenda that accounts for differences in error types but does not address grammatical complexity as a criterion for differentiating error types. She argues that there exists "mounting research evidence that effective error correction—that is selective, prioritized, and clear—can and does help at least some student writers" (Ferris, 1999, p. 4). Ferris maintains that WCF should be used "selectively" to target errors that occur repeatedly in student writing and that "priority" should be given to errors that are most detrimental to meaning. Her reference to WCF being "clear" is not defined according to its capacity to impart declarative knowledge.

Ferris's assertion that WCF needs to be "selective, prioritized, and clear" is relevant to the current study; however, findings here indicate that selective and prioritized use of WCF should be based on whether or not the WCF can provide clear information. That is, WCF should only be used where it can actually be expected to work. Indeed, it is the capacity of WCF to provide "clear" declarative knowledge that determines its effectiveness. Ferris' broad examinations of WCF overall effectiveness do not provide empirical evidence for WCF effectiveness on specific constructions. Because the current study examines different grammatical constructions in detail, construction by construction; we are able to select and prioritize constructions for which WCF will actually work. It does not matter if an error type is more detrimental to meaning if WCF cannot provide clear declarative knowledge that can be practiced, acquired, and used correctly in subsequent writing endeavors.

Ferris (1999) also introduces the concept of treatable versus non-treatable errors. Her distinction between these two error groups is based on whether an error is rule based or idiosyncratic. For Ferris, "rule-based" simply means that rules for a construction can be taught, learned, or looked up in a grammar book (i.e., There is a set of systematic rules.). Based on Ferris's definition of treatable versus non-treatable errors, all grammatical errors and error types would be considered treatable because all are rule based. Non-treatable errors, then, are what she describes as idiosyncratic (i.e., there are no systematic rules to learn). She argues that rule-based systems should be treatable, claiming success in her own classrooms when using WCF to correct errors like subject-verb agreement, run-ons and comma splices, missing articles, and verb form errors (p. 6).

On the other hand, she argues that idiosyncratic errors including "lexical errors and problems with sentence structure…" which include "missing words, unnecessary words, and word order problems" should not be as amenable to change with the use of WCF (p. 6). She states:

> Most systems advocated in writing textbooks and editing handbooks seem based on the assumption that "one size fits all" and that marking "WC" (for word choice) is the same as marking "vt" (for verb tense) ignoring the important fact that for the latter type of error, there are rules to consult, while the former type of error has none. (p. 6)

Ferris's commentary on treatable versus non-treatable errors is relevant to the current research study's findings, which indicate that some errors are more amenable to change than others. However, results of the current study do not correspond to the rule-based/idiosyncratic distinction asserted by Ferris. According to Ferris, the following errors that are examined in this study should be treatable: verb tense, subject-verb agreement, article usage, plural/singular, and sentence fragments. On the other hand, she would consider word order and word usage to be idiosyncratic and non-treatable. Findings from the current study do not confirm Ferris's hypothesis. Instead, WCF effectiveness and treatability appear to depend on varying levels of grammatical complexity.

Finally, the tendency in the academic literature to discuss and reinterpret previous WCF studies (Truscott, 1999; Ferris, 2004; Guenette, 2007; Bitchener & Knoch, 2008) has prompted complaints such as Bruton's (2010) criticism:

…the debate has become "tedious", "sterile" and "academic"… the debate is

tedious because the same points are reiterated; it is sterile because most of the

research central to the argumentation against correction remains the same… with

the numerous recognized flaws and repeated appeals for improved research which

has not been forthcoming; and it is academic in the sense that it does not really

have much relevance for most mainstream L2 writing contexts or practices… (p.

491)

The current research study attempts to address the shortcomings of WCF research by (1)

providing a new research direction for WCF based on a linguistic analysis of WCF

effectiveness, (2) improving methodological controls and utilizing statistical analyses of

data, and (3) examining WCF in such a way that findings are applicable to both SLA and

L2 pedagogy.

Narrow Studies of WCF Effectiveness with the English Article System

Some recent studies of WCF do include control groups and methodologically

sound designs, but they are quite narrowly focused in their coverage of WCF. These

types of studies represent the research agendas of SLA researchers, as opposed to L2-

writing researchers. SLA and L2-writing researchers start by asking very different

questions. SLA researchers want to discover if WCF can facilitate long-term acquisition

of a particular linguistic feature, whereas L2-writing researchers focus on "the question

of whether WCF helps student writers to improve in the overall effectiveness of their

texts and to develop as more successful writers" (Ferris, 2010, p. 188). While SLA

research studies may be more methodologically rigorous than L2-writing studies, the applicability of L2-writing studies to real classroom environments may be more relevant to teachers in the field.

The narrowly focused "article studies" (Sheen Y., 2007; Bitchener, 2008; Bitchener & Knoch, 2008; Bitchener & Knoch, 2010; Ellis R., Sheen Y., Murakami, & Takashima, 2008; Sheen Y., 2010; Shintani & Ellis R., 2013) report that WCF is effective in making students more accurate in two functional uses of the English article system: (1) the use of the indefinite article when mentioning something for the first time and (2) the use of the definite article when the same thing is mentioned subsequently. These studies are examples of focused WCF (focus on one error type only) as opposed to unfocused WCF (focus on a number and variety of writing errors). Generally speaking, experimental groups in all these studies outperform control groups in correcting these two functional uses of English articles over time.

However, this narrow focus on a subset of one linguistic domain demonstrates only that a small area of the English article system is amenable to change. These studies can neither confirm nor refute whether WCF can aid in the acquisition of the English article system in its entirety. As a result, one must question their general applicability to L2-writing classrooms where teachers would like to increase student accuracy for the whole article system and a number of different types of errors. Moreover, an examination of Butler's (2002) comparison of Japanese L2-English students' meta-linguistic knowledge of the English article system and their ability to actually use English articles correctly demonstrates that this type of WCF can in fact be harmful to L2-writing

students. That is, the sub-rule being taught will inevitably run counter to English input and will confuse L2-English students' usage of English articles.

There are multiple exceptions to even this relatively simple "first mention/subsequent mention rule" for article usage. According to Clark (1977), a phenomenon known as "bridging" allows for the first mention of a noun to be definite when it is related to another noun that has already been established as definite:

I saw a house. I walked into the house and entered the kitchen.

The first mention of *house* takes *a*, but the next takes *the*. However, this does not explain how *kitchen* takes *the* with the first mention. Here, because the house is already mentioned in the discourse, interlocutors can bridge their understanding of *the kitchen* as the kitchen in the aforementioned house. Consequently, these studies neither confirm nor reject WCF effectiveness in helping students to acquire real and complete grammatical systems. Moreover, this is precisely this type of grammar instruction that Truscott would consider to be harmful.

It is noteworthy that these researchers do not take on the entire article system. Also, findings of WCF effectiveness on article usage are often mixed with results demonstrating either that only a piece of the system shows increased accuracy or that no improvements occur over the elicitation period. Tentatively, this supports one of Truscott's claims about the ineffectiveness of WCF: Some linguistic domains and the underlying linguistic features of said domains are too complex to be explicitly taught and learned.

41

Bitchener et. al. (2005) finds improvement in the use of the definite article only, and Ferris et. al. (2000) find that students who receive WCF on article usage actually regress in their abilities to produce articles correctly over the elicitation period. Ferris & Roberts (2001) find gains in the accurate use of articles overall compared to a control group, but this study is not longitudinal (only 2 weeks between the first and final draft) and only examines WCF effectiveness as it relates to first and final drafts of the same writing sample, not WCF as a teaching/learning tool that aids in long-term acquisition.

Generally speaking, the literature on the topic of WCF effectiveness and article usage indicates that WCF is either ineffective or only effective in improving parts of the article system, not the article system as a whole. Results of the current research study confirm the ineffectiveness of WCF on increasing article usage accuracy, at least with the indirect type of WCF (correction symbol "art") used in this study.

Different WCF Types and Their Effectiveness

Many studies of WCF have tested the effects of different types of WCF on students' abilities to increase accuracy over time. A brief discussion of the two main types of WCF, *indirect* and *direct,* is relevant to the current research study because results indicate that an SAT analysis of their relative effectiveness could be fruitful.

For both indirect and direct WCF, there are subtypes. Indirect, the type used in the current research study, provides either a correction symbol or some kind of hint about the error that needs correction. For example, errors are circled or the number of errors in a line of text is written next to the line indicating how many errors need to be corrected in

42

that line. Students then have to decipher the symbols or find the errors in order to make

the correction in a subsequent draft. Direct feedback, on the other hand, simply provides

the corrected form of the error that students must make in a subsequent draft. This

correction may or may not include a meta-linguistic explanation of the error correction in

the margins near the correction.

Bitchener and Knoch (2010) explain the theoretical arguments used in support of

indirect and direct feedback as follows:

Those supporting indirect feedback suggest that this approach is best because it

invites L2 writers to engage in guided learning and problem solving and, as a

result, promotes the type of reflection on existing knowledge that is more likely to

foster long-term acquisition and written accuracy. Those more in favor of direct

feedback suggest that it is more helpful to writers because it (1) reduces the type

of confusion that they may experience if they fail to understand or remember the

feedback they have been given (for example, the meaning of error codes used by

teachers); (2) provides them with information to help them resolve more complex

errors (for example, syntactic structure and idiomatic usage); (3) offers more

explicit feedback on hypotheses that may have been made; and (4) is more

immediate (pp. 209-210).

Bitchener and Knoch also comment, "It may be the case that what is most effective is

determined by the goals and proficiency levels of the L2 writers" (p. 210). Generally

speaking, it is assumed that direct feedback may be more effective for low proficiency

students while indirect feedback is more effective for high proficiency learners. Again,

findings here indicate that proficiency should not be the determining factor for which type of WCF is employed. Instead, relative effectiveness should be based on the amount of declarative knowledge imparted by each type of WCF. While this is alluded to in Bitchener and Knock's comments above (e.g., "offers more explicit feedback on hypotheses that have been made"), no attempt has been made to test these theories empirically until now.

There are a number of studies that focus on the relative advantages of different types of corrective feedback. For those studies focusing on the use of direct versus indirect feedback, Lalande (1982) reports that indirect was more effective than direct in increasing participants' accuracy, whereas Semke (1984) reports no difference between the feedback types. Chandler (2003) reports that both indirect and direct feedback had positive effects on increasing accuracy with no major differences between the two types. In studies that compare different types of indirect feedback (i.e. correction symbols, circled errors, number of errors written at the end of a line of text), Robb et al. (1986) and Ferris and Roberts (2001) report no differences between the types under investigation.

Some studies examine the effects of different types of direct feedback by varying which groups receive which types of feedback. In Bitchener & Knoch (2010), there are four groups: Group 1 receives direct WCF (the error is simply corrected for the student), written meta-linguistic explanation, and oral meta-linguistic explanation; group 2 receives direct corrective feedback and written meta-linguistic explanation; group 3 receives direct corrective feedback only; and group 4, the control group, received no WCF. Bitchener (2008), Bitchener & Knoch (2009; 2010), and Ellis R. et. al. (2008)

found no difference between the three groups receiving differentiated feedback but report that the treatment groups all outperform the control group. Sheen (2007) did not find differences between groups that received only direct feedback and those that received written meta-linguistic explanation in the post-test immediately following the study. However, the written meta-linguistic explanation group did significantly outperform the direct feedback only group in a delayed post-test, which lends tentative support to the effectiveness of meta-linguistic explanation for long-term acquisition of linguistic items.

A brief review of the literature assessing different types of WCF demonstrates that experiments have been conducted in much the same way examinations of WCF/no-WCF groups have been conducted. That is, researchers simply compare the different types and make broad claims that one is better than another or that there is no difference between the types. They provide no hypothesis for *why* one might be more effective than another, and they conduct no experiments to test such a hypothesis. The current research study can neither confirm nor refute claims for which type is more effective. It was not designed to do so. However, results here indicate that examining the different types according to the principles of SAT could help in our understanding of *why* one is more effective than others.

Indeed, different WCF types could be assessed according to their capacity to impart declarative knowledge, which leads to meaningful practice, which leads to proceduralization of knowledge. Direct feedback that simply provides the correct verb tense in given situations would certainly be more informative (i.e., provides more declarative knowledge and clearer feedback) than indirect feedback that hints that an

error has been made but does not provide the correct revision. Under the indirect

paradigm, L2 students are left to guess at the correct revision, and they cannot be sure

they have made the correct revision once it has been made.

However, in the case of binary option constructions, indirect WCF symbols

highlight an error for the L2 learner and also provide the correct answer by default (i.e.,

the other option), as with correction of a true/false test that has been previously marked

by a teacher. In these cases, indirect feedback behaves very much like direct feedback,

though proponents of indirect WCF might argue that having students decipher the symbol

and provide the correct option themselves promotes the type of engagement in the

practice procedure that better facilitates proceduralization. That is, simply providing the

correct revision, as direct feedback does, might reduce engagement and problem solving

within the practice procedure, and this could lead to less evidence of proceduralization.

Of course, this could be tested empirically by examining indirect WCF's effectiveness on

binary structures and direct WCF's effectiveness on those same structures. On the other

hand, non-binary grammatical paradigms can be likened to the challenge of correcting a

multiple choice test when a teacher has simply marked the answer wrong. Here, the L2

student is required to choose from among many options to make the correction and

cannot be sure he or she has made the correct revision. Direct feedback should provide

more useful declarative knowledge than indirect for non-binary forms (e.g., "b" is the

correct answer), and the relative effectiveness of the two types of feedback could be

assessed based on their capacity to provide different amounts of declarative knowledge.

This topic is discussed in more detail under the discussion sections in Chapter 8.

Wagner 2013

A brief review of results from Wagner (2013) is pertinent at this time because findings from that study provide the impetus for the current research study's design and SAT analysis of WCF effectiveness. Because Wagner (2013) examines WCF effectiveness broadly and narrowly (i.e., its overall effectiveness and its effectiveness on individual forms), we were able to determine not only its overall effectiveness but also its effectiveness on individual constructions. In a comparison of experimental and control groups, findings demonstrate overall effectiveness of WCF for reducing a combination of all error types (Corresponding ANOVA P Value = .004 with a statistical significance level of < .05) and effectiveness for the subcategories linguistic/grammatical errors (Corresponding ANOVA P Value = .045), word usage errors (Corresponding ANOVA P Value = .044), and mechanical errors (Corresponding ANOVA P Value = .086 with a marginal statistical significance level of < .1). These findings demonstrate a general usefulness of WCF for making students more accurate writers.

However, a more detailed analysis of the data provides us with a clearer picture of WCF utility, especially regarding WCF effectiveness for increasing students' grammatical accuracy. Of the four linguistic errors types (subject-verb agreement, singular/plural, verb tense, and article usage) examined as a part of Wagner (2013), subject-verb agreement is most amenable to positive change. Despite the relatively low sample size (n=33) and the relatively low number of errors (We are now only looking at total errors and averages for one error type.), a statistically significant difference was

discovered between the control and experimental groups for improving subject-verb agreement (Corresponding ANOVA P Value = .008). Also, experimental groups outperformed control groups for reducing singular/plural errors across the elicitation tasks (Corresponding ANOVA P Value = .05).

Conversely, there are no statistically significant differences between experimental and control groups for verb tense (Corresponding ANOVA P Value = .790) or for articles (Corresponding ANOVA P Value = .127). Experimental group participants received feedback on all error types, interacted with the symbols for all four types, and made corrections to all four types in subsequent drafts of their writing samples. However, they do not demonstrate increased grammatical competence compared to control group participants for verb tense or article usage.

The discovery of WCF's variable effectiveness in Wagner (2013) provides the foundation for the research of this dissertation. Results indicate that the delineation between treatable and non-treatable errors is not based on whether or not constructions correspond to the rule-based/idiosyncratic distinction asserted in Ferris (1999). Instead, grammatical complexity appears to delineate for which constructions WCF will be effective and for which it will not.

CHAPTER THREE: RESEARCH DESIGN OVERVIEW

The current research of WCF examines WCF in real classroom settings (i.e., an L2-English classroom and an L2-Spanish classroom) and focuses on a variety of error types in order to replicate how WCF is actually utilized by L2-writing instructors. Data is analyzed to determine not only WCF's overall effectiveness but also its effectiveness on individual constructions. This research clarifies considerably why and how WCF either succeeds or fails in teaching different grammatical constructions.

The current study demonstrates precisely what "effective" means regarding the use of WCF. Here "effectiveness" is measured in two ways: (1) by how well WCF imparts clear declarative knowledge of grammatical constructions to L2 learners (L2-English WCF Uptake Study) and (2) by how well it aids in the transition to proceduralization of this knowledge via practice (L2-Spanish WCF Study). After all, it is the underlying assumption of WCF practitioners that WCF will transmit knowledge that can be practiced via the correction process and eventually become proceduralized. Findings from the current research study indicates that while the practice procedure remains constant across all the grammatical categories (i.e., the correction of errors in subsequent drafts of the same writing sample), indirect WCF effectiveness in providing declarative knowledge varies based on the complexity of the grammatical constructions under examination. WCF effectiveness then is measured based on a linguistic analysis of

the grammatical constructions themselves, and such analyses provide a clearer prediction

as to when WCF will be highly effective and when it will be useless or virtually useless.

Research Question and Hypothesis

This study addresses the following research questions and hypothesis, repeated

below for the reader's convenience:

> **RQ 1:** Does WCF and the correction process aid in the transition from
> declarative to proceduralized knowledge of specific grammatical constructions in
> the writing modality? And if so, for which constructions is this transition
> observed?

> **RQ 2**: Given the answers to RQ1, is there a grammatical complexity threshold
> beyond which WCF will no longer be operative in supplying the declarative
> knowledge necessary for L2-writing students to practice and eventually
> proceduralize?

It would be possible to imagine that a complexity threshold might be sensitive to many

diverse factors. However, as an introductory research direction based on preliminary

findings from Wagner (2013), this study tests the following initial hypothesis:

> **Hypothesis**. WCF is only effective in imparting declarative knowledge thereby
> increasing linguistic accuracy (proceduralization) for grammatical constructions
> that include a binary choice of options for correct usage (binary vs. n-ary ($n > 2$)).

Two Studies Overview

L2-English WCF Uptake Study. The proposed research contains two separate

research studies. One takes place in an L2-English classroom (The L2-English WCF

Uptake Study) and focuses more specifically on WCF symbols' capacities to impart clear

declarative knowledge for a variety of grammatical forms. In this study, participants'

behaviors with making corrections based on indirect WCF symbols between rough and final drafts are examined. This type of examination demonstrates how well WCF does or does not impart declarative knowledge.

For example, when a student repeatedly receives the symbol *s-v* above verbs in rough drafts and then makes the corrections by either adding or removing an "s" from the verbs, providing WCF is tantamount to providing the student with the correct answer (i.e., WCF communicates in effect *This is wrong. The other option is right*, as there is only one other option). When students encounter present tense verbs in new writing samples, they thus have a binary choice to make: *supply the agreement marker or not based on the composition of the subject*.

The symbol *num* (*number* error- for singular/plural errors) plausibly works in precisely the same manner. Where students receive the symbol, the *other* option is correct, as no other options exist in English for correct usage. Declarative knowledge is clearly imparted (i.e., *Add "s" for plural nouns. Do not add "s" for singular nouns.*), and the practice of correcting the errors helps to lead to proceduralization of the grammatical rule as demonstrated by statistically significant increases in singular/plural accuracy in subsequent writing samples.

However, the added linguistic complexity of verb tense and article usage appears to have rendered WCF ineffective. The symbols *vt* and *art* draw L2-writing students' attentions to the grammatical errors and to the grammatical categories the symbols represent. Students attempt to make the corrections to these errors; however, these two symbols do not impart enough declarative knowledge in the form of simple repair

strategy that students can practice and eventually proceduralize. For example, when an L2-writing student of English receives the symbol *vt*, it is only possible to guess at the correction from among the many options of tense marking (i.e., twelve tenses constructed with a large variety of linguistic materials) and from among the variety of semantic uses of tenses in English. Upon encountering the production of a verb in a subsequent piece of writing, the student has not been taught a deterministic rule to follow in producing the verb in the correct tense.

The same is true of articles. When the student receives the symbol *art*, it is clear that an error has been made, but it is not possible to glean a rule from the WCF alone. English articles are notoriously difficult for L2-English learners to acquire. Whereas advanced English students can often acquire fluency and a great deal of accuracy across other linguistic domains, articles often remain recalcitrant to change even with L2-English learners at the most advanced proficiency levels (Kharma, 1981; Yamada & Matsuura , 1982; Butler, 2002). This is due in large part to the fact that English articles are not comprised of one-to-one form and meaning associations (Anderson R. , 1984). Consequently, a correction symbol does not impart the declarative knowledge necessary to elucidate the abstractness and opacity of the English article system. Regardless, the hypothesis for the current study is that WCF fails to impart a learnable repair strategy even before the opacity of the article system (or other opaque/complex systems) is taken into account due to more fundamental linguistic considerations, which is that the grammatical pattern includes more than a binary choice for correct usage.

In the case of the English article system, there are three choices: the definite article *the*, the indefinite article *a/an*, and the null article Ø. The abstractness and opacity of the article system's form-meaning relationships aside (WCF cannot impart knowledge of these relationships either), L2-English writing students learn only that they are using articles incorrectly when they receive the *art* correction symbol, but not what correction to make. Although they have a 50% chance of guessing the right alternative between the remaining two options available in the article system, these students cannot be sure they have chosen correctly and have not learned a specific rule to apply with articles when approaching noun phrases in subsequent writing assignments. Arguably, this is precisely the underlying intuition that has led SLA/WCF researchers to narrowly examine the effects of WCF on only one sub rule of the English article system that focuses on a *binary choice* (*a/an* for the first mention of an NP and *the* for subsequent mentions of an NP) (Sheen Y., 2007; Bitchener & Knoch, 2008; Ellis R., Sheen Y., Murakami, & Takashima, 2008; Bitchener & Knoch, 2009; Bitchener & Knoch, 2010) instead of focusing on the acquisition of the article system as a whole.

Based on the current research hypothesis, it was predicted that participants would make corrections based on WCF between first and final drafts of the same writing samples better for subject-verb agreement and singular/plural errors (i.e., the binary option errors) than for verb tense and articles errors (i.e., the non-binary option errors). It was hypothesized that study participants would demonstrate knowledge of the binary option constructions in their corrections to final draft errors while demonstrating no such understanding of the non-binary option constructions.

In addition to the four grammatical categories discussed above, a decision was made to include the grammatical category "prepositions" with the corresponding WCF symbol *pr*. Based on the current research hypothesis, it was predicted that the sheer number of English prepositions and the idiosyncratic nature of preposition usage would also render WCF as an ineffective teaching tool to explicate and impart the rules for correct usage. Prepositions, like English articles, are extremely difficult for L2-English students to acquire and use properly. People can be *in a car*, *in a bus*, or *in a train*; but you can also be *on a bus* or *on a train*. To be *on a car* is to be sitting *on top of a car*. In some cases, a person may *scream at another person*, but a person can also *scream to someone* to do something. An L2-English student who is confronted with the symbol *pr* for a preposition error has an enormous amount of prepositions from which to choose when making the correction in a subsequent draft of the same writing sample. The symbol provides no assistance in selecting the proper preposition, and it does not explain why the preposition provided (or not provided) is wrong. Even a student who makes the correct selection from among the hundreds of English prepositions cannot be sure they have made the right selection. This type of practice, devoid of clear declarative instruction, cannot lead to proceduralization. Regardless, the WCF symbol *pr* is often used to point out students' preposition errors in L2-English writing classrooms. Examining WCF effectiveness on preposition errors provides an additional opportunity to test the binary vs. n-ary (n > 2) hypothesis. Based on the research hypothesis, it was predicted that preposition errors would not be revised correctly between drafts (i.e., unsuccessful uptake) as well as the binary option constructions.

Examining WCF via SAT allows us to understand the reasons why WCF is sometimes effective and sometimes is not. In the cases of subject-verb agreement and singular/plural usage in English, WCF is effective because the symbols *s-v* and *num* not only draw students' attention to the error but also provide the correct answer for how to fix the error (i.e., the other choice of the binary choice). In the cases of verb tense, article usage, and preposition usage in English; useful declarative knowledge cannot be imparted by a symbol pointing out an error. Consequently, there is no rule to practice and no rule to proceduralize. Options abound with correct verb tense, article usage, and preposition usage. In effect, L2-writing instructors are merely pointing out these errors with no guidance on how to correct them. This is probably the basis of Truscott's claim that WCF can actually be harmful to L2-writing students. One can imagine a student becoming frustrated as their errors are continually pointed out and corrected but never reduce in frequency.

As mentioned previously, no argument is being made that proceduralization of grammatical knowledge can *only* occur via overt declarative training. Increased grammatical competence in an L2 can occur as a result of a variety of variables. However, the strength of the current study's design is that it controls for such variables and focuses on WCF symbols' relative effectiveness as the single variable responsible for the success or failure of proceduralization of grammatical rules in the writing modality, as seen in Chapter 4.

L2-Spanish WCF Study. Examining WCF effectiveness for teaching grammatical constructions in a language other than English broadens the applicability of

findings to more than just the L2-English classroom. Also, Spanish provides an

opportunity to test a binary pattern that exhibits idiosyncrasy, which allows for a more in-

depth analysis of the binary vs. n-ary (n > 2) hypothesis. This study examines L2-

Spanish students' error reduction tendencies between control (no WCF provided) and

experimental groups (WCF provided) for new writing samples over an extended period of

time. In this study, we are directly examining which error types are and are not amenable

to positive change (i.e., accuracy increases/proceduralization) via the use of WCF

symbols. Accuracy increases represent how we operationalize the concept of

proceduralization for the purposes of this study.

An investigation of Spanish experimental and control group L2-learners provides

additional opportunities to demonstrate that a binary ceiling exists above which WCF will

no longer be effective. First of all, based on the current study's research hypothesis, it

was predicted that L2-Spanish writers' in the experimental/treatment groups would

significantly increase accuracy on singular/plural errors over time (i.e., proceduralization)

compared to control/no-treatment groups because the singular/plural rule in Spanish is

generally a binary choice like in English (add "s" or do not add "s"). However, it was

also predicted that subject-verb agreement in Spanish would not be amenable to change

with the use of WCF. Spanish, unlike English, is a morphologically rich language and

correct Spanish subject-verb agreement is dependent upon a large number of options for

correct usage:

1. **Yo** habl**o**
   I speak

   Tu habl**as**

You speak

El habl**a**
He speaks

Nosotros habl**amos**
We speak

Vosotros habl**ais**
You (pl) speak

Ellos habl**an**
They speak

Consequently, it was hypothesized that subject-verb agreement errors would not significantly reduce over time for experimental versus control groups. Likewise, the binary vs. n-ary (n > 2) hypothesis predicts that Spanish verb tense and article usage will not show statistically significant improvements because they are similarly complex to English verb tense and article usage (i.e., include more options than a binary choice for correct usage).

Crucially, Spanish provides another binary choice grammatical form that allows us to test the binary vs. n-ary (n > 2) hypothesis: gender marking and agreement. All language items in Spanish noun phrases must agree in gender, and correct usage depends on a binary choice (either masculine-2 or feminine-3):

2. **el**   teléfon**o**   roj**o** fe**o**
   the   telephone red ugly
   the ugly red telephone

3. l**a**   biciclet**a** roj**a** fe**a**
   the bike   red ugly
   the ugly red bike

Of course, there are minor exceptions to this rule. However, it was hypothesized that regular masculine/feminine agreement would significantly improve for experimental groups relative to control groups in the L2-Spanish study despite any outlying irregular usage in the form. According to the binary vs. n-ary (n > 2) hypothesis, the *mf* symbol used to draw students' attention to masculine/feminine errors should be enough to convey the necessary declarative knowledge of the grammatical rules underlying the construction. Also, upon closer examination of occurrences of masculine/feminine errors, it became clear that a distinction between gender selection of individual words and gender agreement across noun phrases had to be made, as seen in more detail below.

The researcher acknowledges that high school L2-English students living in America and high school L2-Spanish students living in America are not directly comparable because the L2-English students' ambient language is English while the L2-Spanish students' ambient language is *not* Spanish. However, the differences between the two types of study participants are irrelevant. The Spanish study stands on its own and focuses only on WCF effectiveness (based on the binary vs. n-ary (n > 2) hypothesis) and the acquisition of Spanish grammatical constructions in the writing modality over a 12-week period. During the elicitation period, L2-Spanish students had equivalent opportunities to receive a variety of input (i.e., reading L2 Spanish and conversation with native speakers of Spanish) for all Spanish grammatical constructions. Where statistically significant gains are made in grammatical competence in the writing modality, these gains can be attributed the use of WCF symbols that convey clear and

useful declarative knowledge. Where gains are not made, the results demonstrate that declarative knowledge was not imparted by said symbols.

Other types of input (e.g., reading and conversations with native speakers) for the L2 under analysis (Spanish) were consistent for all the individual grammatical constructions under analysis. That is, L2-Spanish speakers had as much opportunity to receive a variety of input on Spanish singular/plural as they did on Spanish subject-verb agreement, Spanish verb tense, Spanish article usage, and Spanish masculine/feminine agreement. Although these other forms of input probably influenced participants' abilities to proceduralize grammatical forms, the use of control groups and the statistical analysis of data collected from the WCF study ensures that WCF effects are being measured and not these outside influences.

The intervention of WCF symbols that flags a variety of grammatical errors acts as the only independent variable leading to varied accuracy changes (i.e., the dependent variable) among the grammatical categories. Students received feedback on errors after each of the first three elicitation tasks (rough drafts). They used this feedback to correct the errors in their final drafts (i.e. the practice). This correction practice was constant for all grammatical categories under examination, and only a symbols' capacity to convey declarative knowledge was evaluated via participants' abilities to reduce errors (i.e., proceduralization).

.

CHAPTER FOUR: L2-ENGLISH WCF "UPTAKE" STUDY METHODOLOGY

This chapter describes the methodology used in the L2-English WCF Uptake Study. First, the research design and context are described. Second, the participants, instruction and intervention, and instruments are described. Third, data collection and analysis is explained.

Research Design

This study is an experimental design and examines the differences in participant behaviors when they encounter and attempt to correct errors which have been grouped into one of two categories: (1) binary option constructions (i.e., subject-verb agreement and singular/plural) and (2) non-binary option constructions (i.e., verb tense, articles, and prepositions). This study examines WCF's ability to impart declarative knowledge of simple and complex linguistic constructions based on the binary vs. n-ary (n > 2) hypothesis. Simply put, participants should demonstrate knowledge of binary option constructions and a lack of knowledge of non-binary option constructions.

This study does not necessitate the use of experimental and control groups because we are not examining WCF effectiveness for helping L2 students' transition from declarative to proceduralized knowledge. Instead, this study examines the interaction of L2 students with the WCF symbols themselves in order to determine which

symbols impart sufficient declarative knowledge and which do not. This study examines the beginning of the SAT transition process (WCF and declarative knowledge provided and students attempt to decipher and use said knowledge) as opposed to the process as a whole. It focuses on the dichotomy between binary option constructions and non-binary option constructions and the divergent behaviors of participants attempting to correct these two types of constructions.

Research Context

The investigation for the L2-English WCF Uptake Study was conducted in five English Language Learner (ELL) classrooms at a high school in eastern Kansas. This course emphasizes improving students' academic writing skills by focusing on the writing process, organization in writing and effective argumentation, and grammatical accuracy. All five sections of the class were taught by the researcher in the same way. Throughout this yearlong course, students learned to write science reports, journals, fiction and nonfiction stories, academic essays, business letters and resumes, and summaries.

Participants' English proficiency levels were similar and determined their placement in the class. While proficiency levels of students vary, all the students were competent enough to meet the linguistic demands of the writing course. Each class was composed of no more than twelve students from culturally and linguistically diverse backgrounds. They were required to take this writing course to strengthen their writing skills for mainstream English classes at the high school and to ensure they maintained or

improved their state assessment scores on the writing section of the Kansas English Language Proficiency Assessment (KELPA), as seen in more detail below. The curriculum for the course focuses on improvement in writing through multiple drafts instead of only evaluating a grammatically perfect writing sample turned in as a final draft.

An Overview

The L2-English WCF Uptake Study analyzes the "uptake" of declarative knowledge *during* the correction process, helps to answer research questions 1 and 2, and confirms the binary vs. n-ary (n > 2) research hypothesis. Under this current "uptake study," participants' rough and final drafts for the three elicitation tasks are compared to scrutinize where students do and do not make the correct revisions on their final drafts based on the symbols provided in the rough drafts.

Here, the term "uptake" refers to ELL students' abilities to make the *correct revision* between the rough and final drafts. This study examines corrections between the rough and final drafts and provides invaluable information regarding the students' understanding of the symbols and the students' abilities to make the correct revisions between drafts of the same writing samples based on the WCF provided. Thus, the primary purpose of the L2-English WCF Uptake Study is to examine WCF symbols' abilities to impart clear knowledge for varied grammatical constructions.

For example, if a student receives the symbol *s-v* in a handwritten rough draft and either supplies or does not supply the agreement marker correctly in the final draft (Word

document with error alerts and notifications turned off) according to the symbol provided, this is considered successful "uptake" of declarative knowledge and successful practice. It is precisely this type of practice that can facilitate a transition from declarative to proceduralized knowledge. However, if a student receives the symbol *vt* on a rough draft and alters the verb incorrectly in the final draft or does not make any change to the verb in the final draft, this is considered unsuccessful uptake of declarative knowledge and such practice (i.e., the act of recognizing that an error has been flagged with a symbol in the rough draft and making the correction in the final draft based on the symbol provided) is not only ineffective but could also be considered detrimental to proceduralization of said knowledge. Findings help to reinforce when WCF is effective at imparting declarative knowledge and when it is not. Before running this L2-English study, it was hypothesized that the binary choice constructions would be revised correctly between drafts (i.e., uptake is successful and practice is meaningful) significantly better than constructions that include more than a binary choice for correct usage.

This L2-English uptake study focuses on the WCF symbols more directly. It examines symbols' decipherability, not students' proceduralization of knowledge over time. Therefore, this study does not require the use of control groups. The focus of the study is to examine the differences between correction behaviors of participants dealing with binary option corrections and non-binary option corrections. All L2-English writing students received WCF symbols for errors in the grammatical constructions under analysis. Then rough and final drafts of the same writing samples were compared to examine where symbols provided enough declarative knowledge for successful uptake

63

(i.e., correct revisions of flagged errors in the final draft) and where symbols did not provide enough knowledge for successful uptake (i.e., incorrect or no revisions of flagged errors in the final draft).

Participants and Proficiency Levels

This study originally consisted of 47 ELL students from a high school in eastern Kansas. At the time of the study, participants' composite English proficiency scores were between an upper intermediate score of 3 and a high proficiency score of 4 according to the Kansas English Language Proficiency Assessment (KELPA). Over the course of the six-week elicitation and research period, 14 students were released from the study because they moved out of the area, attended too few classes to complete the elicitation tasks in a timely manner, or changed their class schedule at the high school. Consequently, data for the ELL study was collected from a total of 33 ELL students, 18 females and 15 males.

In Kansas, ELL students' initial proficiency and their annual proficiency progress are measured using KELPA. KELPA assesses students reading, writing, speaking, and listening skills as well as their social and academic language skills. Composite scores (overall proficiency) are based on a combination of students' four individual language skill scores. Students' overall proficiencies are ranked from 0 to 4. Students who score a composite score of 4 two years in a row are exited from ELL services, and they attend only mainstream high school classes thereafter.

Participants are predominately from Central and South America, the majority from Mexico (26). However students from a number of other countries are represented including: Peru (1), Kenya (1), South Korea (1), Ethiopia (3), Congo (1), and Iraq (1). These 9th through 12th grade students' ages range from 14 to 18. Their grade levels are determined by a combination of previous education in their home countries, English proficiency based on KELPA scores, and the number of mainstream classes they have already attended and passed in the United States. Participants have had between three and five years of instruction in the United States, and they all have composite KELPA scores between 3.0 and 4.0 and KELPA writing scores between 3.0 and 4.0 (See Table 4.1 for the KELPA Writing Rubric).

Table 4.1: KELPA Writing Rubric

Rating	Vocabulary	Sentence Fluency	Grammar	Mechanics	Organization and Development
4	Descriptive and vivid vocabulary is specific and enhances meaning.	The writer uses a variety of sentence structures, including compound and/or complex sentences, correctly and effectively.	Verb tenses and subject/verb agreement are varied, appropriate, and correct. Any errors that are present do not impede meaning.	There are few or no errors in punctuation, capitalization, and spelling. The writer uses indentation and alignment for paragraph format.	There is a clear main idea with details that are relevant, specific, and appropriate to the prompt. Supporting ideas are well organized with a clear beginning, middle, and end. Transitions are used correctly.
3	Vocabulary is adequate but not expressive; words are more general than specific. A few usage errors may be present but do not impede communication.	Sentences have some variety and may attempt compound structures. A few errors may be present, but do not impede communication.	Writing may contain verb tense and subject/verb agreement errors. Errors do not interfere with meaning.	Some errors in punctuation, capitalization, and spelling are present.	The paper is focused on a single idea or event related to the prompt, but may provide only simple supporting details as well as some extraneous details. Writing has a recognizable flow from start to finish, but there are few or repetitive transitional signals.
2	Vocabulary is limited with usage errors that may impede communication.	Sentences consist of basic structural patterns and may contain errors that may impede meaning. Sentence construction may be awkward.	Writing contains frequent errors in verb tenses and subject/verb agreement. Errors might interfere with meaning.	Frequent errors in punctuation, capitalization, and spelling.	There is a main idea or focus, but details are unrelated.
1	Vocabulary is limited to isolated words and/or phrases.	There are no complete sentences.	Writing contains consistent errors in verb tense that impede understanding.	Consistent errors in punctuation, capitalization, and spelling.	Writing only consists of repetitive simple sentences, fragments, or isolated words.
0	There is no response or the response is mostly in a language other than English.	There is no response or the response is mostly in a language other than English.	There is no response or the response is mostly in a language other than English.	There is no response or the response is mostly in a language other than English.	There is no response or the response is mostly in a language other than English.

Participants have a mix of ELL support classes and mainstream high school classes. They were placed in this particular writing class specifically for writing support based on their KELPA scores from the previous year. Those students with a score of 4 scored 4 on writing for the first time the previous year, and the goal of the class was to ensure they scored 4 on writing again during the current school year. For those students who scored 3 on writing the previous year, the goal was to improve their writing, so they could score a 4 for the first time during the current school year. Though proficiency KELPA proficiency scores are not a perfect measure of students' writing ability, it does provide a relatively reliable assessment for placement. Participants could write well enough to produce longer works in different rhetorical forms but still struggled with grammatical errors and corrections in their writing samples. Table 2 provides a breakdown of information on individual participants:

Table 4.2: Participant Information

Participant # and Gender	Class Section	Grade	Age	Home Country	L1	KELPA Writing Score
1 m	1	10	15	Mexico	Spanish	3
2 f	1	10	15	Mexico	Spanish	4
3 m	1	10	16	Mexico	Spanish	4
4 m	1	10	16	Mexico	Spanish	3
5 m	1	9	15	Mexico	Spanish	4
6 m	1	9	15	Mexico	Spanish	4
7 f	2	10	16	Mexico	Spanish	4
8 m	2	10	16	Congo	French	3
9 f	2	11	17	Mexico	Spanish	3
10 m	3	11	17	Mexico	Spanish	3

11 m	3	11	17	Mexico	Spanish	3
12 m	3	9	14	Mexico	Spanish	4
13 m	3	9	14	Mexico	Spanish	3
14 f	3	9	14	Mexico	Spanish	4
15 m	3	10	16	Mexico	Spanish	3
16 f	3	9	14	Mexico	Spanish	4
17 f	3	12	18	Peru	Spanish	3
18 f	3	9	15	Mexico	Spanish	4
19 m	4	9	14	Mexico	Spanish	4
20 f	4	11	17	Mexico	Spanish	4
21 m	4	9	15	Mexico	Spanish	4
22 f	4	11	17	Kenya	Kiswahili	4
23 f	4	11	17	Iraq	Arabic	4
24 f	4	12	18	Mexico	Spanish	3
25 f	4	11	18	Mexico	Spanish	4
26 m	4	9	14	Ethiopia	Amharic	3
27 m	4	9	14	Mexico	Spanish	4
28 m	4	9	15	Mexico	Spanish	3
29 f	5	10	15	Mexico	Spanish	3
30 f	5	10	16	Mexico	Spanish	4
31 f	5	12	19	Ethiopia	Amharic	3
32 f	5	11	18	Ethiopia	Amharic	4
33 f	5	12	18	Korea	Korean	4

The sampling method used in the current study is a convenience sampling, which is a non-probability sampling method. In convenience sampling, the selection of participants is based on availability or accessibility. I am an ELL teacher with access to L2-English writing students for over twelve years. Therefore, students were available to assist in my examination of WCF effectiveness and more specifically my research questions and research hypothesis. All participants volunteered for the study and filled out either assent or consent forms before the study began. For minor participants, parents also completed a consent form before the study began. These assent and consent forms can be seen in the appendix of this study.

Instruction

In order to control for the effects of WCF, a concerted effort was made to hold other variables constant. All participants received four and a half hours of instructional time per week. Over the six-week treatment and observation period for this study, all instructional content having to do with writing was the same for all participants among all five sections of the class with a focus on writing strong essays in different rhetorical forms. Besides these essays, students wrote journal entries, which were based on the same prompts in each class and were never corrected for grammar, punctuation, or mechanics. WCF was reserved for so-called academic essays only. WCF had not been used prior to the start of the study.

Participants followed a five-step writing process for completing each elicitation essay under examination in this study. Each class completed two paragraphs prior to the start of the study using the writing process as modeled by the teacher. During these modeling sessions, students were instructed on higher order aspects of writing paragraphs. They learned how to organize ideas into main ideas (topic sentences), major supporting details, minor supporting details, and conclusion sentences. The five steps of the writing process used for the current study include: (1) prewriting, (2) organizing, (3) drafting, (4) revising and editing, and (5) publishing. All participants received WCF on their papers as a part of the revising and editing step (step 4) of the writing process. Participants were then asked to publish their essays using the correction symbols and the correction symbol key to make the proper corrections.

Targeted Errors and Treatment

Twelve errors types were targeted and tabulated from the participants' writing samples. These errors types can be further categorized by their intended instructional purpose as seen in Table 3:

Table 4.3: Error Categories and Types

Error Category	Error Type
Linguistic (5)	Subject Verb Agreement / Verb Tense / Singular, Plural / Article Usage / Preposition Usage
Word Usage (2)	Wrong Word (used for homonyms only) / Word Form (parts of speech)
Mechanics (5)	Capitalization / Punctuation / Run-ons / Fragments / Spelling

These error types, and the corrections symbols used to represent these error types, were chosen for the current study because they are representative of the most common errors that students make in their writing and of the errors that ELL teachers at this particular high school focus on when correcting students' writing. These three categories of error types were also chosen to determine if participants could attend to and acquire the linguistic forms (subject-verb agreement, verb tense, singular/plural, article usage, and preposition usage) while attending to other types of errors simultaneously.

Before the actual study of WCF "uptake" began, students were provided with the following correction symbol key:

Table 4.4: Correction Symbol Key

Symbol	Symbol Meaning	Example
s-v	subject-verb agreement	He <u>love</u> his wife.
vt	verb tense	I <u>will be</u> in class yesterday.
num	number- singular/plural	I have many <u>dog</u> at home.
cap	capitalization	<u>doug</u> went to NYC on <u>tuesday</u>.
ro	run on	He likes to go to the store, and he likes to go to the beach, but he does not like to go to the park when it is hot because there is no place to cool down, and his sister also hates the park in the summer time, but their mother loves to walk in the park any time of the year, so she goes there every day to exercise.
frag	fragment	If you like pizza.
wf	word form	He completed his <u>apply</u> for college last night.
ww	wrong word	<u>There</u> house is very beautiful.
P	Punctuation (. ? ! , ; : ' ")	What is he doing__

sp	spelling	My <u>freind</u> is <u>funy</u>.
pr	preposition	He goes <u>at</u> the store.
art	wrong or missing article (a, an, the)	My teacher is <u>a</u> best one in ____ school

Participants made corrections to the examples under the "example" section of the key and the corrections were discussed as a class to determine whether or not they were correct. Also, small mini-lessons were presented on each error type in the two weeks preceding the actual study. Finally, students were given a worksheet that included writing with all of the different error types. These errors were marked with the WCF symbols, and students were tasked with correcting the errors. When they were finished practicing on their own, we looked at the errors and corrections as a class. All of this was done to ensure students had a good working knowledge of WCF and the symbols and to ensure they had clear expectations for how they should correct final drafts using WCF in future writing assignments. After this "introduction period," students began to produce the three authentic writing samples that are being used as a part of this study.

Elicitation Tasks

Thirty-three students wrote three, five-paragraph essays on topics of their choice for this "uptake" study, which took place over a six-week period. Students finished handwriting each essay over the course of two, fifty-minute class periods. When students

finished writing each essay, they turned it in, WCF would be provided by the researcher, and they would rewrite it in the following class using their correction symbol keys and making corrections to errors marked with WCF. The final draft of each elicitation task was completed on Microsoft Word with error indications turned off. They did this three separate times over the six-week period. Participants completed the first elicitation task (rough draft) and correction task (final draft) during week one, the second during week three, and the final during weeks five and six.

Data Collection Procedures and Analysis

When all the writing samples were complete, rough and final drafts were compared to determine precisely when students revised flagged errors correctly (i.e., successful uptake of declarative knowledge and effective practice that can potentially lead to proceduralization) and when they did not revise them correctly or did not revise them at all (i.e., unsuccessful uptake of declarative knowledge and ineffective practice not leading to proceduralization). Comparing the error types binary versus non-binary with participants revision behaviors (correct revisions versus incorrect or no attempt at revisions) demonstrates that declarative knowledge was imparted by WCF and used with the binary construction revisions but was not imparted by WCF and used with the non-binary construction revisions. The number of each error type was calculated for each student on each of the three essays in the following order: subject-verb agreement, singular/plural, verb tense, articles, and prepositions.

For example, the first rough draft essay for each student was first examined to find precisely where subject-verb agreement errors had been made. When one was discovered in the rough draft, the precise location of the subject-verb error was examined in the final draft of the typed writing sample. Subject-verb errors were highlighted in red, and the number of these errors revised correctly was tabulated to determine the rate or percentage of times students revised errors correctly for each category type. So, if a participant had two subject-verb errors in a particular rough draft of an essay but only corrected one of them in the final draft of the same essay, this would be a 50% correction rate for subject-verb errors for that particular participant in that particular essay. The same was done for each error type across all three elicitation tasks. Singular/plural instances were highlighted in blue in final drafts, verb tense in green, articles in brown, and prepositions in purple. Using the color codes helped to make quick comparisons between rough and final drafts much easier.

Once the color-coding and tabulation of errors versus errors corrected was complete, data was entered into an excel spreadsheet under elicitation tasks 1, 2, and 3 with subfields under each task for each of the five grammatical categories. Each category was further divided into the total number of errors for each error type and the number of times it was revised correctly. Once this was complete, comparisons could be made among the different error types and correction rates.

Rates of correct revisions were compared between binary and non-binary constructions, and this analysis demonstrated that students behavior with binary constructions differed greatly compared to their behavior with non-binary constructions.

While they revised nearly 80% of binary option constructions correctly, they only corrected about 50% of the non-binary option constructions correctly. Once it became clear that the non-binary option constructions were corrected correctly only about 50% of the time, more detailed analyses of verb tense, article, and preposition correction behavior was conducted to examine participant behavior with these more complex forms. This analysis focused more specifically on participants' correction behavior when encountering a *vt*, *art*, or *pr* symbol. Tabulations were made examining students errors that were not revised correctly. These were coded for "correct correction," "no correction attempt," and "new error." In some cases, this detailed examination discovered other interesting participant behaviors seen in more detail below.

For verb tense, it became clear during coding that simple present and simple past corrections were being revised more accurately than other tenses. In most cases, these errors could be corrected by either adding or deleting "ed." This is, of course, a binary choice, and participants appeared to be dealing with such error corrections more accurately than corrections that needed more complex revisions (e.g., simple present to present perfect corrections). Consequently, the number of "ed errors" was tabulated and compared to the number of times these errors were revised correctly. The same was done for " > ed" errors. Also, when correct revisions were made, notes were taken on the type of error and how the revision was made. Indeed, these notes helped to distinguish participant behaviors with "ed" and " > ed" errors, as seen below.

For articles, "correct correction," "no correction attempt," and "new error" tabulations were made as well. Notes were taken on how correct revisions were made in

order to examine precisely when participants might show an understanding of article usage from among the many uses. This was instructive. Almost all the correct revisions made to article errors were made from an incorrect use of the null article to the correct use of either the definite or indefinite article.

The same procedure was used for preposition error and revision calculations. Notes taken on correct revisions of preposition errors demonstrate that participants excelled mostly on correctly repairing in/on errors. Again, this is a binary choice. For example, participants corrected things like "in Thanksgiving day" to "on Thanksgiving day" and "on the U.S." to "in the U.S.

Descriptive statistics and statistical analysis for all the data can be seen in Chapter 5 below. Statistical analysis of the binary constructions is compared to data from the n-ary (n > 2) constructions using a Z-Test and the following statistical hypotheses and rules:

> **Null Hypothesis**: There is no difference between binary and n-ary (n > 2) grammatical constructions on the effective use of WCF in providing declarative knowledge to students that can be practiced correctly. In other words, the means of the two samples are the same and the WCF symbols are equally likely to be either effective or ineffective in providing clear declarative knowledge that can be practiced meaningfully.

> **Alternative Hypothesis**: There is a significant difference in the mean. The mean of binary constructions' data is significantly higher than n-ary (n > 2) constructions' data.

> **Rules for Rejecting the Null Hypothesis**: If P-Value is less than 0.05 (95% confidence level of the results), then the difference is significant; otherwise, accept the null hypothesis.

This tabulation and analysis of errors allow us to determine which symbols provide clear declarative knowledge for effective practice and which do not. According to the binary vs. n-ary (n > 2) hypothesis, subject-verb agreement and singular plural WCF symbols (*s-v* and *num* respectively) should provide the necessary declarative knowledge to ensure students make the correct revisions to these binary pieces of grammar. Of course, there is nothing intrinsically superior about these symbols when compared to their counterparts: verb tense (*vt*), articles (*art*), and preposition (*pr*) symbols. The *s-v* and *num* symbols are simply pointing out errors for grammatical constructions that can be corrected according to a binary choice. The other symbols point out errors for grammatical constructions with multiple forms and a large variety of linguistic materials and form-meaning relationships.

CHAPTER FIVE: L2-ENGLISH WCF "UPTAKE" STUDY RESULTS

Statistical Analysis

Data comparing binary vs. n-ary (n >2) grammatical constructions support the alternative statistical hypothesis and the current research hypothesis. Data from the L2-English WCF Uptake Study allows for a rejection of the statistical null hypothesis and an acceptance of the statistical alternative hypothesis. Results are highly significant (P-Value less than 0.0001 at a confidence level of 0.05). Binary constructions were revised correctly a majority of the time while constructions with more than two correct usage options were corrected only about 50% of the time, were corrected incorrectly, or were not corrected at all. WCF appears to provide the necessary declarative knowledge for subject-verb agreement and singular/plural in English. This, in turn, allows for meaningful, correct practice of these forms (i.e., correcting errors), and this leads to the proceduralization of these binary rule types as confirmed by the L2-Spanish WCF Study. Verb tense, article, and preposition corrections were characterized by additional errors or no corrections at all.

Descriptive Statistics

General findings. Participants (n=33) make a total of 792 grammatical errors in first drafts of their three essays over the three elicitation tasks. Of these errors, they

revise 501 correctly (63%) and 291 incorrectly or not at all (37%) in final drafts, as seen

in Figure 5.1. Initially, in Figures 5.1 through 5.3, incorrect revisions and no attempt at

revisions are not disaggregated. We begin by examining total errors and simply looking

at which errors were revised correctly in the final drafts and which ones were not revised

correctly (either as a result of making a new error or not attempting a revision) in the

final draft. As we begin to investigate participant correction behavior with the individual

grammatical constructions verb tense, articles, and prepositions; we take a more detailed

look at the data by splitting the "incorrect revision behavior" and "no attempt at revision

behavior" (Figures 5.4, 5.7, and 5.8).

It is not necessary to make this distinction with subject-verb agreement and

singular/plural behaviors. Participants revise these error types correctly a vast majority

of the time, and they do not make any new errors in the final drafts on these

constructions. That is, the very few unrevised errors for subject-verb agreement and

singular/plural are all "no attempt at revision." Based on the overwhelming number of

correct revisions of these two error types, we can assume participants simply overlook the

occurrences of these errors and forget to revise them in the final draft. Still, it is

instructive to recognize that the binary option construction revisions do not include new

errors in subsequent drafts. This, undoubtedly is linked to the forms' inherent simplicity.
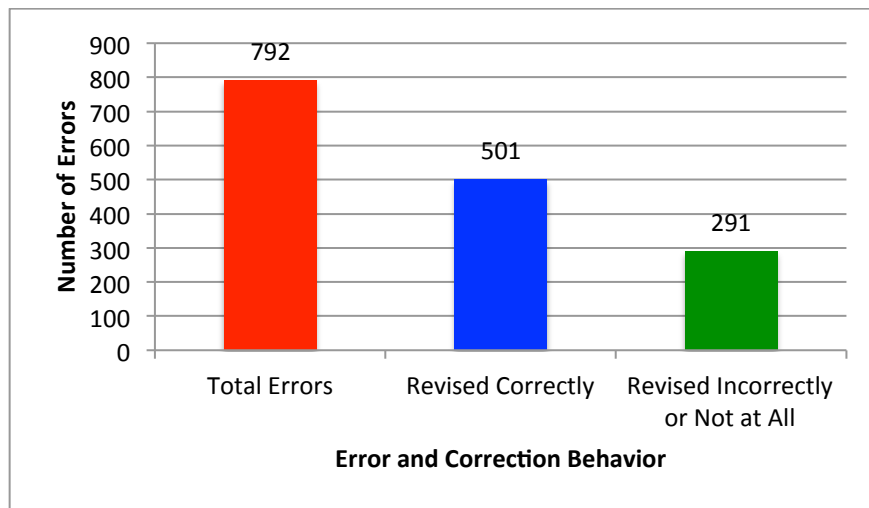
Figure 5.1: Total Errors, Total Correct Revisions, Total Incorrect Revisions

Participants make a total of 99 subject-verb errors in the first drafts and revised

85 of them correctly (86%) in the final drafts. They make a total of 187 singular/plural

errors and revise 170 of them correctly (91%) in final drafts. They make a total of 271

verb tense errors but only revise 137 of these errors correctly (51%) in the final drafts.

They make a total of 115 article errors and only revise 58 of these errors correctly (50%)

in final drafts. Finally, participants make 120 preposition errors and only revise 51 of

these errors correctly (43%) in the final drafts. Findings for each error type can be seen

in Figure 5.2 below. Again, the "Total Corrected Incorrectly" in green in Figure 5.2 does

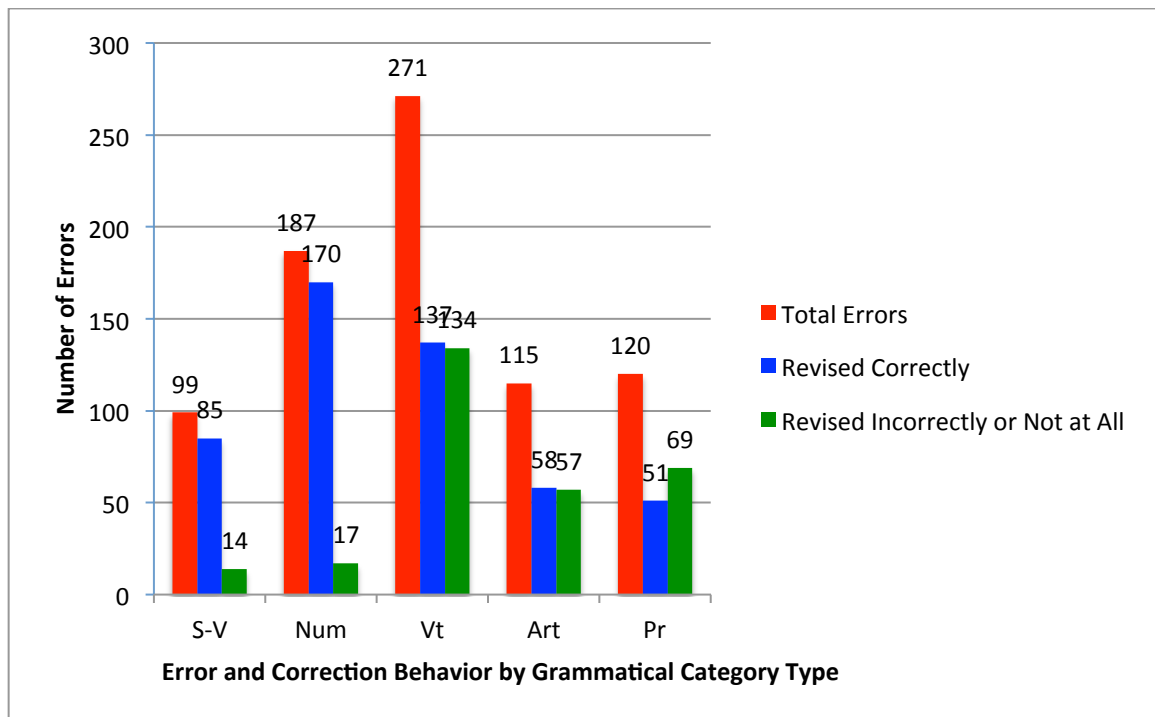not yet split errors that were revised incorrectly from those that were not even attempted.

Figure 5.2: Total Errors by Grammatical Category, Correct Revisions by Category, and Incorrect Revisions by Category

These findings are precisely what we should expect to discover based on the binary vs. n-ary (n > 2) hypothesis. The binary option constructions (i.e., subject-verb agreement and singular/plural) errors are revised correctly a vast majority of the time (86% and 91% respectively). This demonstrates that the symbols *s-v* and *num* provide enough declarative knowledge for these particular constructions to be corrected in subsequent drafts of the same writing sample. As a result, practice is meaningful and effective. L2-writing students clearly understand the symbols, the meaning of the symbols, and the grammatical rules represented by the symbols. This type of correction practice can lead to proceduralization of knowledge.

Conversely, the non-binary option construction errors (i.e., verb tense, articles, and prepositions) are revised correctly only half or less than half of the time (51%, 50%, and 43% respectively). These correction percentages demonstrate that students are either guessing at how to repair errors or are simply giving up on repairing them because they have no decipherable knowledge to apply during the correction/practice procedure. Clearly, this type of practice is not useful, and L2 learners have no hope of proceduralizing knowledge that has not been imparted or learned.

A total of 286 binary grammar errors were made (i.e., subject-verb agreement errors and singular/plural errors). Of these errors, 225 were revised correctly (79%). A total of 506 non-binary grammar errors were made (i.e., verb tense, articles, prepositions), and only 246 were revised correctly (49%), as seen in Figure 5.3.
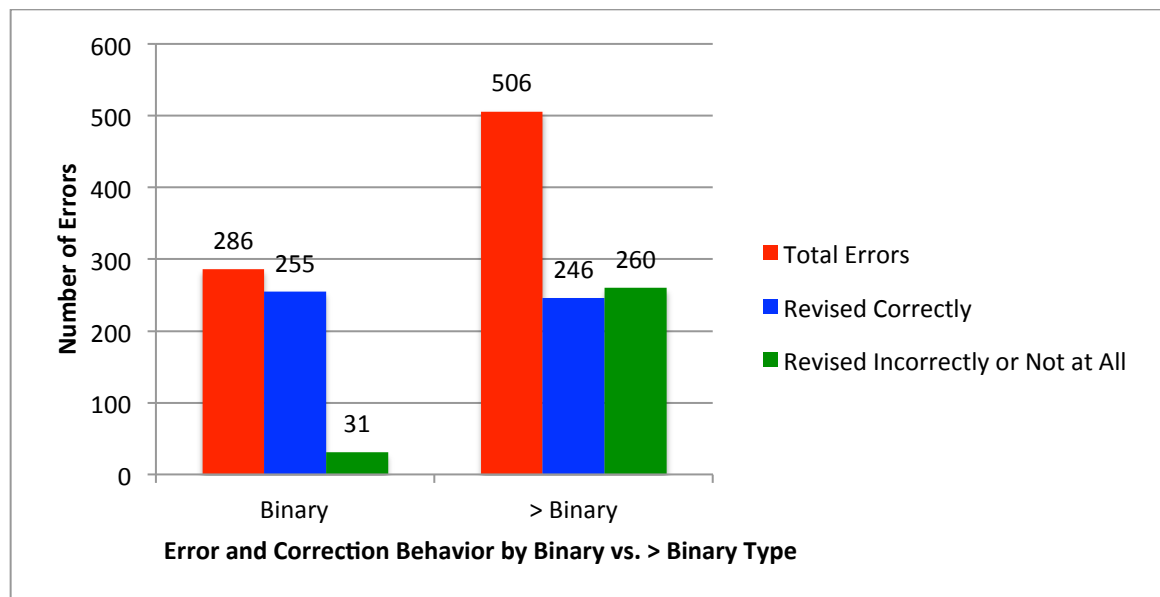

Figure 5.3: Binary versus > Binary Total Errors, Correct Revisions, and Incorrect Revisions

Again, this shows a vast difference between the behaviors of participants when confronted with the two types of grammatical constructions. They correct subject-verb agreement errors and singular/plural errors at a rate that confirms knowledge of the constructions' underlying rules.   Conversely, their inability to correct the non-binary option constructions confirms that WCF does not impart enough declarative knowledge for L2 learners to practice and eventually proceduralize.

Generally speaking, when participants are confronted with either an *s-v* or *num* symbol in their rough drafts, they either add the agreement/plural marker or delete it to make the correct revision in the final draft. In most cases and for both of these binary error types, "s" morphology needed to be *added* to the verb or to the noun respectively to correct the error. Deletions of unnecessary "s" morphology occur only a handful of times.

Verb Tense Findings. Participant correction behavior in final drafts when confronted with *vt*, *art*, and *pr* symbols in rough drafts demonstrates a lack of declarative knowledge and/or confusion. A detailed examination of their attempts to make corrections to non-binary option grammatical categories reveals similar correction behaviors within each of these grammatical categories.

Here we make a closer examination of precisely how participants attempt to correct or do not attempt to correct verb tense errors. This detailed examination of their correct revisions, their non-attempts, and their new errors provides insight into just how ineffective WCF is at imparting the necessary declarative knowledge of English verb

tense usage to L2-English learners. Of the 271 verb tense errors made by the 33 participants, 134 are not revised correctly in final drafts (49%). Of those 134, no attempt to correct is made on 99 (74%) of the errors. Of those 134, new errors are made on 35 (26%) attempts to correct original errors, as seen in Figure 5.4.
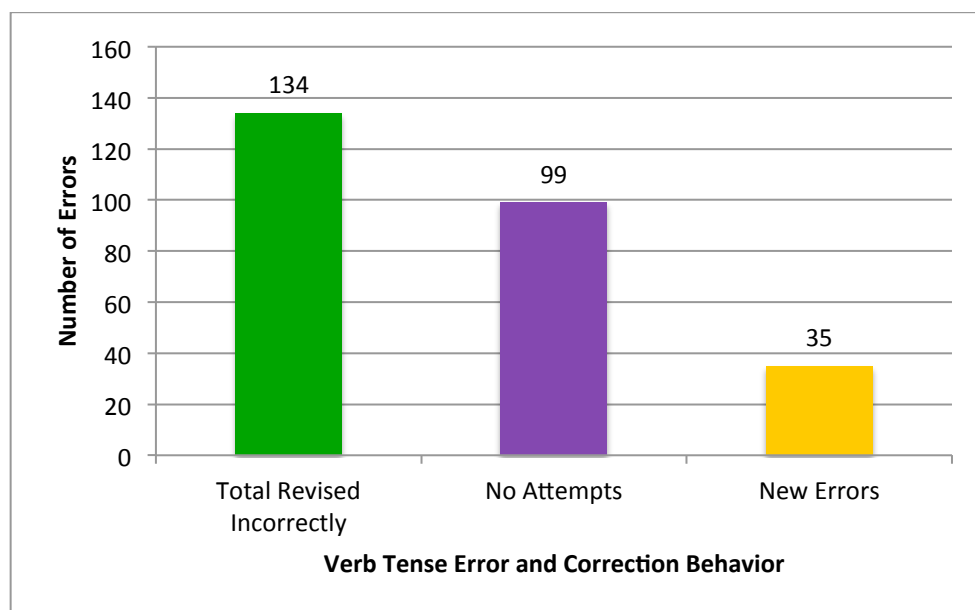
Figure 5.4: Verb Tense Total Revised Incorrectly, No Attempts at Revision, and New Errors

This demonstrates that declarative knowledge of the English verb tense system has not been established in participants via WCF (i.e., the pointing out of an error). In a vast majority of encounters with verb tense errors and WCF, participants simply do not try to correct the error in the final draft. It is clear that they have no knowledge of how to proceed during the correction process. Their behavior demonstrates that they do not have

declarative knowledge of the grammatical category and that the practice procedure is therefore ineffective.

Looking more closely at the verb tense errors that were revised correctly (137 of 271 or 51%), we see that 87 (64%) of the errors can be correctly amended by either adding or removing *ed*. Of those 87 "add ed remove ed" errors, participants revise 57 correctly (66%), as seen in Figure 5.5.
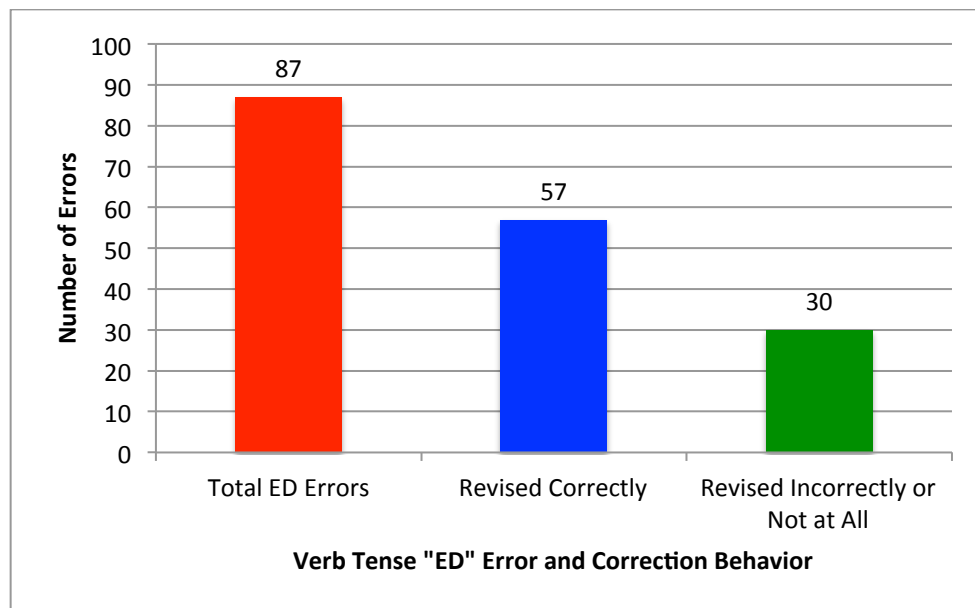

Figure 5.5: Total Verb Tense "ED" Errors, Correct Revisions, and Incorrect Revisions

This demonstrates a least a marginal improvement regarding participants' abilities to deal with simple present tense and simple past tense errors relative to verb tense grammar as a whole. More important, correcting such errors is a binary choice (i.e., add "ed" or delete

"ed"); however, proper English verb tense usage depends on far more than a declarative or proceduralized knowledge of simple present and simple past tenses. Still, it appears as if participants were looking for a decipherable repair strategy that they could remember and execute properly without confusion over other options for correct usage.

Indeed, when participants encounter errors that require knowledge of more complex verb tenses (i.e., present perfect, past perfect, etc.) in order to make proper corrections, they rarely correct the errors correctly. There are 22 instances of errors that required such knowledge, and only 5 (23%) are revised correctly in final drafts, as seen in Figure 5.6.
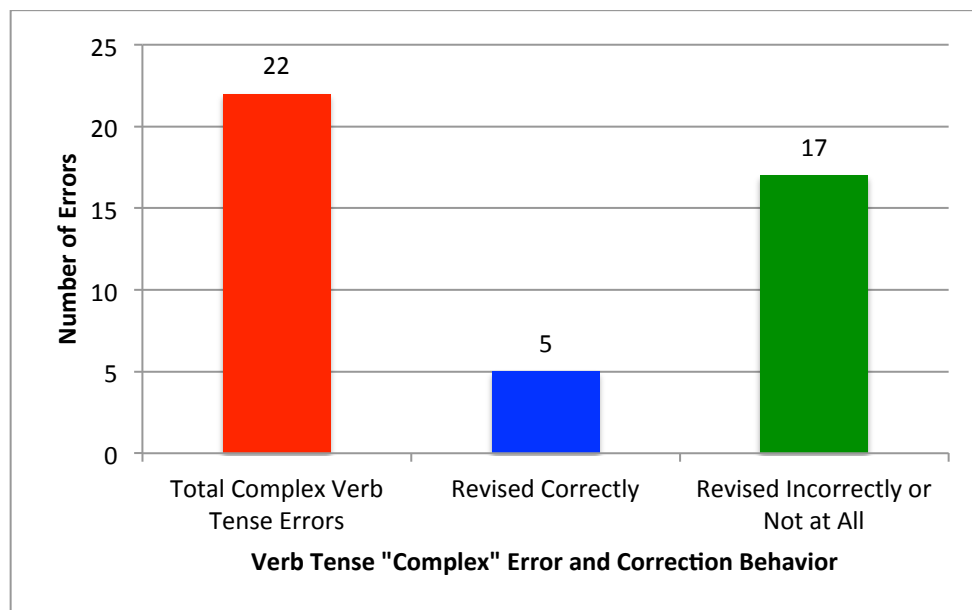
Figure 5.6: Total Complex Verb Tense Errors, Correct Revisions, and Incorrect Revisions

Participant 4 changes "have gave" to "have given." Participant 22 changes "have living" to "have been living." Participant 23 correctly changes "has offered" to "offered." Participant 28 changes "is" to "has always been." This particular example represents the only correct correction from such a simple form to such a complex form across the entire data set. Participant 29 changes "are gone" to "would be gone."

Otherwise, participant behavior when encountering complex verb tense corrections demonstrates a lack of knowledge or confusion. Participants generally do not attempt to correct these errors. When they do, they make new errors. Participant 25 changes "had accomplished" to "have accomplish." In the same sentence, she changes "won" to "win," but she should have changed the construction to "have won." Participant 31 changed "had teached" to "had teach," possibly in an attempt to apply the "add/delete ed" rule. The same participant made no attempt to fix "never been," "never seen," and "I been shy" in the same writing sample. A close examination of participant behavior with verb tenses in the correction process demonstrates that WCF is not an effective instructional tool for expounding the underlying complexities of the English verb tense system.

Article Findings. Participant uptake behavior for articles is similar to their uptake behavior for verb tense. Of the 115 article errors across all three elicitation tasks, participants revise only 58 of the errors correctly (about 50%), as seen in Figure 5.2 above. Of these 58 correct article corrections, 55 of the corrections were made by either changing a null article to the indefinite article (no distinction was made between *a* and *an*) or to the definite article. In only two instances did participants correctly change an

86

indefinite or definite article to the null article (in these 2 instances, both were originally the indefinite article). This could explain why correct article revisions only occur 50% of the time. It appears that participants are guessing between the indefinite or definite article when they encounter an empty space before a noun phrase that requires a correction. It could be that they are simply flipping a coin between the indefinite and definite article when a correction needs to be made. They do not realize that the null article could also be used for correct usage in English noun phrases.

Of the 57 times they fail to make the correct article revisions, they make no correction at all in the final draft 37 times (65%) and make new errors on 20 attempted corrections (35%), as seen in Figure 5.7.
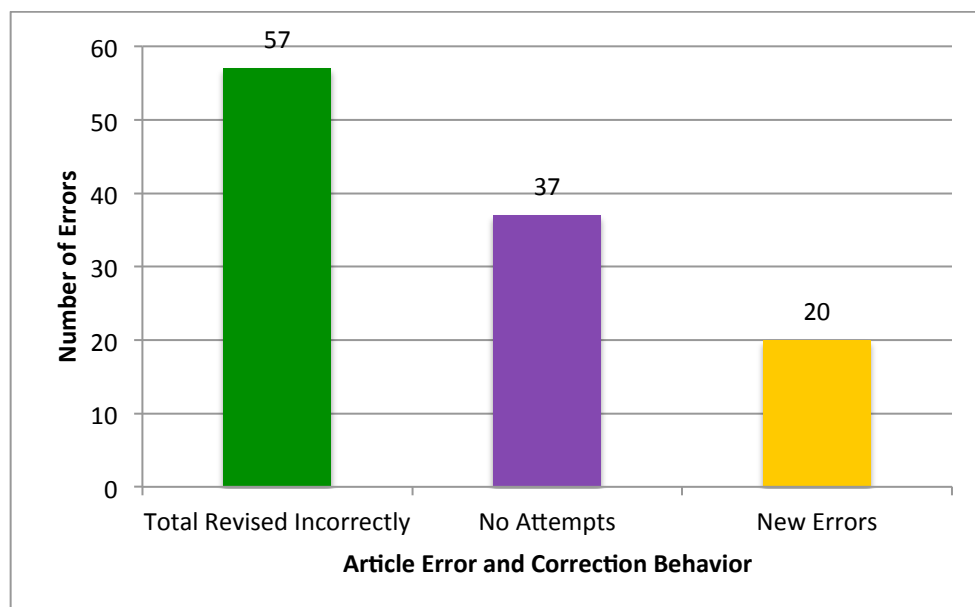

Figure 5.7: Articles Total Revised Incorrectly, No Attempts at Revision, and New Errors

Once again, this demonstrates that participants who are capable of making correct revisions to subject-verb and singular/plural errors do not have the knowledge necessary to even attempt corrections to non-binary option errors like article errors. When attempts are made, participants make new errors, which also demonstrates a lack of knowledge about how to correctly revise and/or use the English article system.

        Preposition Findings. Similar to their verb tense and article uptake behaviors, participants' corrections of prepositions demonstrate a lack of declarative knowledge based on the WCF provided. Of the 120 preposition errors, only 51 (about 43%) are revised correctly in final drafts, and 69 (about 57%) are not revised, or they are revised incorrectly. Of the 69 errors not corrected in final drafts, participants make no attempt to correct 38 (55%) of the preposition errors. They make 31 (45%) new errors in attempting to correct errors in final drafts, as seen in Figure 5.8.
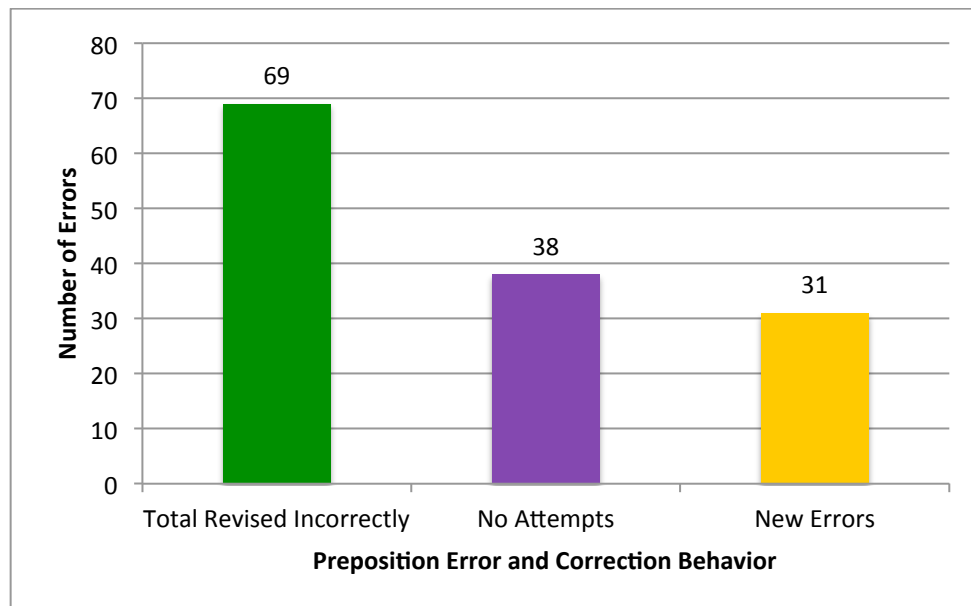
Figure 5.8: Preposition Total Revised Incorrectly, Total No Attempts at Revision, and New Errors

This is one major difference between participants' preposition and verb tense and article correction behavior.   Participants make no attempt to correct 74% of verb tense errors and 65% of article errors. At least with prepositions, they attempt to make a correction about 50% of the time. However, these attempts do not demonstrate a better understanding of preposition usage for participants who correct only 43% of preposition errors correctly relative to 51% and 50% for verb tense and article errors respectively.

Findings, Research Questions, and the Research Hypothesis

Overall, findings from the L2-English WCF Uptake Study provide the following answers to the studies research questions:

RQ 1: Does WCF and the correction process aid in the transition from declarative to proceduralized knowledge of specific grammatical constructions in

the writing modality? And if so, for which constructions is this transition observed?

Although findings from this study do not examine the transition from declarative knowledge to proceduralized knowledge directly (it was not designed to do so), it confirms a potential breakdown in the SAT transition process from the beginning. The L2-English WCF Uptake Study was designed to examine the interaction between study participants and the WCF symbols themselves, not participant proceduralization over time. This examination permits us to verify which symbols provide sufficient declarative knowledge and which do not. WCF can aid in the transition process if it provides clear declarative knowledge of grammatical constructions. The transmission of clear declarative knowledge is a prerequisite for meaningful practice. WCF cannot aid in the transition process if it falters in providing the necessary knowledge from the outset, which takes us to the second research question.

> **RQ 2**: Given the answers to RQ1, is there a grammatical complexity threshold beyond which WCF will no longer be operative in supplying the declarative knowledge necessary for L2-writing students to practice and eventually proceduralize?

Findings confirm a grammatical complexity threshold beyond which WCF will no longer be operative. Binary option constructions are revised a vast majority of the time, while non-binary option constructions are either not revised in final drafts or are revised incorrectly about half the time. Therefore, WCF is not successful in providing enough declarative knowledge for non-binary option constructions, and practice for said constructions is not constructive toward a proceduralization of this insufficient knowledge.

Therefore, the binary vs. n-ary (n > 2) is upheld at least in part:

Hypothesis. WCF is only effective in imparting declarative knowledge thereby increasing linguistic accuracy (proceduralization) for grammatical constructions that include a binary choice of options for correct usage (binary vs. n-ary (n > 2)).

Although the L2-English Uptake Study alone cannot confirm proceduralization of grammatical rules, findings confirm that WCF is only effective in providing sufficient declarative knowledge for subject-verb agreement and singular/plural constructions. These findings taken in connection with findings from the L2-Spanish WCF Study below demonstrate that a binary complexity barrier exists for the successful implementation of WCF. Further discussion of these results in combination with results from the L2-Spanish WCF Study can be in the discussion section in Chapter 8 below.

CHAPTER SIX: L2-SPANISH WCF STUDY

This chapter describes the methodology used in the L2-Spanish WCF Study. First, the research design and context are described. Second, the participants, instruction and intervention, and instruments are described. Third, data collection and analysis is explained.

Research Design

This study is an experimental design and examines the differences between control group participant and experimental group participant proceduralization behaviors for grammatical constructions which have been grouped into one of two categories: (1) binary option constructions (i.e., singular/plural and masculine/feminine agreement) and (2) non-binary option constructions (i.e., subject-verb agreement, verb tense, and articles). This study examines WCF's ability to impart declarative knowledge of simple and complex linguistic constructions based on the binary vs. n-ary $(n > 2)$ hypothesis. Based on the research hypothesis, participants are predicted to demonstrate proceduralized knowledge of binary option constructions and a lack of proceduralized knowledge of non-binary option constructions over the elicitation period.

Research Context

The investigation for the L2-Spanish WCF study was conducted in two Spanish 5 classrooms at a high school in eastern Kansas. This course places emphasis on improving students' speaking, listening, reading and writing skills in Spanish. Both classes were taught by the same bilingual (Spanish/English) language teacher in precisely the same way throughout the year and over the elicitation period, and classes were conducted entirely in Spanish. Although the classes were firmly based in meaning-based methodologies and communicative language activities, there was some focus on grammar.

The focus on grammar, however, was mostly implicit with very little overt grammar instruction or practice. For example, students were required to write about and discuss "things they did last week" using the simple past tense in Spanish to correctly express their activities. During this activity, the teacher provided recasts and oral feedback but made no overt explanations of the simple past in Spanish. From what the teacher explained, most of the L2-Spanish curriculum and her lesson plans were developed in a similar way. The teacher did comment on the explicit nature of WCF being used in the study as, "not something the students would be used to."

Students generally had been taking Spanish for either three or four years prior to being enrolled into this Spanish 5 class. Although their proficiencies vary, they all have

enough Spanish skills to fulfill the writing requirements for both regular class assignments and for the elicitation tasks used as a part of this study.

An Overview

This Spanish WCF study was specifically designed both to confirm and augment to the results of Wagner (2013) using a different L2 with similar and dissimilar grammatical constructions and to test the validity of the binary vs. n-ary (n > 2) hypothesis. As mentioned previously, Spanish subject-verb agreement, article usage, and verb tense constructions are grouped as non-binary option constructions while singular/plural and masculine/feminine agreement are grouped as binary option constructions according to the research hypothesis.

It was predicted that Spanish singular/plural errors and masculine/feminine agreement errors would reduce for experimental group participants significantly more than control group participants across the elicitation tasks and over the elicitation period (i.e., accuracy increases and proceduralization of grammatical rules), while the non-binary option constructions would not show significant differences in reductions between the two groups. Closer examination of the data necessitated a more refined analysis of the binary vs. non-binary distinction, as seen below.

Participants and Proficiency Levels

This study originally consisted of 53 students from a high school in eastern Kansas. These participants are all advanced Spanish speakers and attend one of two

Spanish 5 classes at the high school. Over the course of the 12-week research period, 13

participants were released from the study because they were taking the class as heritage

language learners (i.e., students who already speak some Spanish because it is spoken at

home or who speak Spanish as their first language but are taking the class to learn more

about their native language), attended too few classes to complete the elicitation tasks in

a timely manner, or changed their class schedule at the high school. As a result, the

current study includes a total of 40 participants, 21 in the experimental group and 19 in

the control group.

Students were placed into Spanish 5 as a result of having fulfilled prerequisite

requirements for placement (i.e., Passing Spanish 1, 2, 3, and 4) or passing an exam

demonstrating the knowledge necessary to be successful in the class. Both Spanish 5

classes were conducted entirely in Spanish by a bilingual Spanish teacher. All

participants speak English as their native language and all were between the ages of 16

and 18. See Table 13 for information on regarding experimental versus control groups.

Table 6.1: Participant and Group Information

Section	Number of Participants	Native Language	Proficiency	Experimental / Control
1	21	English	ADVANCED	E- WCF
2	19	English	ADVANCED	C- No WCF

Instruction

95

In order to control for the effects of WCF, a concerted effort was made to hold other variables between the two classes constant. Both classes received four and a half hours of instructional time per week. Over the twelve-week treatment and observation period for this study, all instructional content having to do with writing and grammar was the same for both classes (i.e., the same lesson plans were delivered). WCF was only used for the writing samples used as a part of this study.

Grammar, punctuation, and word usage mini-lessons were delivered in precisely the same way in the two classes, except that the experimental groups received a correction symbol key that linked specific correction symbols to the grammatical principles, word usage issues, and mechanical problems being covered in class. Before the elicitation tasks were undertaken, instruction was given to all participants (control and experimental) on each concept that each correction symbol represents. The researcher and the Spanish teacher delivered these lessons over the course of two fifty-minute class periods. The lessons consisted of two error correction worksheets. Control groups attempted to correct the errors with no guidance. When they were done, the researcher and teacher went over the corrections as a class, and students either confirmed they corrected the errors correctly or changed their incorrect corrections to match the agreed upon correction. On the other hand, experimental groups were provided with a correction symbol key (seen below), and feedback was provided using symbols from this correction symbol key. The researcher and teacher followed the same procedure regarding class review and correction of errors. These lessons were concluded before the 12-week

elicitation task period began so as not to affect student error rates during the elicitation period.

Only the experimental groups received WCF on their writing samples. These experimental groups were then asked to publish their paragraphs using the correction symbols and the correction symbol key to make the proper corrections. The control group was simply asked to "look for errors and try to correct them" before the publishing stage of the writing process.

Targeted Errors and Treatment

Twelve errors types were targeted and tabulated from the participants' writing samples. These errors types can be further categorized by their intended instructional purpose as seen in Table 14.

Table 6.2: Error Categories and Types

Error Category	Error Type
Linguistic (5)	Subject Verb Agreement / Verb Tense / Singular, Plural / Article Usage / Masculine, Feminine Agreement
Word Usage (2)	Wrong Word (used for homonyms only) / Word Form (parts of speech)
Mechanics (5)	Capitalization / Punctuation / Run-ons / Fragments / Spelling

These error types, and the corrections symbols used to represent these error types, were chosen for the current study because they are representative of the most common errors

that students make in their writing and of the errors that SLL teachers at this particular high school focus on when correcting students' writing. These three categories of error types were also chosen to determine if participants could attend to and acquire the linguistic forms (subject-verb agreement, verb tense, singular/plural, article usage, and masculine/feminine agreement) while attending to other types of errors simultaneously.

Although both experimental and control groups received explicit instruction on grammar, word usage, and mechanics represented by the correction symbols; only the experimental group received WCF and a correction symbol key (Table 15) to help them decipher the meanings of symbols during the publishing stage of the writing process. This key includes the correction symbols, the meanings of the correction symbols, and examples of when each correction symbol would be used.

Table 6.3: SLL Correction Symbol Key

Symbol	Symbol Meaning	Example
s-v	subject-verb agreement	Jose am<u>o</u> su esposa.
vt	verb tense	Yo <u>ire'</u> a la escuela ayer.
num	number- singular/plural	Yo tengo muchos <u>perro</u> en me casa.
cap	capitalization	<u>jose</u> va a <u>nuevo york</u> ayer.

ro	run on	<u>Yo quiero ir a la escuela, y yo quiero ir a la tienda de mi madre, o nosotros podemos ir a ver una pelicula, pero yo no quiero ir a Nuevo York porque no me gusta ciudades grandes, y yo no tengo dinero bastante para ir a Nuevo York, y tu tampoco tienes dinero bastante ir Nuevo York.</u>
frag	fragment	<u>Si te gusta pizza</u>.
wf	word form	El completo' su <u>aplicar</u> anoche.
P	Punctuation (. ? ! , ; : ' ")	__Que esta haciendo__
sp	spelling	Mi <u>amijo</u> es muy contento.
mf	masculine/feminine	Mi plum<u>a</u> roj<u>o</u> esta rot<u>o</u>
art	wrong or missing article (el, los, la, las, un, una, unas)	La profesora es la mejor de __ escuela.

When experimental group participants were provided with the correction symbol key, they followed the teacher's lead in writing the correction symbol in for each of the errors and correcting the errors under the "example" section of the key. This key was provided to experimental participants only and was used during the publishing step of the writing process for each of the elicitation tasks. The control group never received or saw this key.

Elicitation Tasks

Each participant completed four paragraphs (min. 200 words each) over the twelve-week period at weeks one, four, eight, and twelve. Although students were required to write a descriptive paragraph for week 1, a process/analysis paragraph for week 2, a compare/contrast paragraph for week 3, and a cause and effect paragraph for week 4; they were allowed to choose their own topics. Students were provided with short examples of each type of paragraph in English by the researcher in order to ensure that they understood how to write each type of piece. As most of the students were upper classmen at the high school, they were already familiar with all four of these rhetorical forms and had written each type in their English classes. There is a general tendency for error rates to increase in the fourth elicitation task, especially among control group participants, but it is difficult to determine if that is result of the rhetorical form.

After completing the prewriting and organizing steps of the writing process during one class period for each writing sample, students were given fifty minutes in a the following class period to complete the rough drafts of each paragraph. Correction symbols were added by researcher in collaboration with a native Spanish-speaking Spanish teacher at the high school to the experimental groups' paragraphs between classes. Both experimental group and control group classes published their final drafts in the following class using Microsoft Word with error notifications turned off.

Data Collection Procedures and Analysis

After students had completed all four elicitation tasks, data was collected from all four of each of the L2-Spanish participant's writing samples and input into an Excel

100

spreadsheet. Data collection consisted of counting each of the twelve error types, one at a time, in each of the four paragraphs in order to determine how many of each type of error occurred in each writing sample. For example, singular/plural errors were calculated first. Every occurrence of the WCF symbol *num* for each participant was counted in writing samples 1, 2, 3, and 4 and was documented accordingly in the Excel spreadsheet. The same collection procedure was used for every error type.

Tabulating masculine/feminine agreement errors presented a unique challenge that the researcher did not anticipate. Initially, all masculine/feminine agreement errors were tallied as incorrect if participants did not choose the correct gender of the noun included in the noun phrase. This was the result of simply counting all the *mf* symbols that were supplied within the noun phrases by the researcher and L2-Spanish teacher in order to (1) have the participants change the noun gender from the original gender supplied to the correct gender and (2) make other things in the noun phrase match that gender, as seen in number 1 below:

<p style="text-align:center;">mf   mf   mf</p>
<p style="text-align:center;">1. *el plumo rojo</p>
<p style="text-align:center;">the   pen   red</p>
<p style="text-align:center;">the   red   pen</p>

In Spanish, the word for pen is a feminine noun and should be written as *la pluma*, and adjectives modifying nouns in a noun phrase must agree in both number and gender. Therefore, agreement dictates that the correct usage for the noun phrase in number 1 is *la pluma roja*, as "pen" in Spanish is a feminine noun and "a" endings generally designate feminine.

It became clear that data representing student knowledge of particular nouns' genders and their knowledge of agreement within noun phrases had to be teased apart. Knowing a noun's gender and knowing how make linguistic materials within a noun phrase agree are two separate types of knowledge. Therefore, two coding categories were created in order to tabulate the two types of knowledge and to determine if the *mf* symbol had an effect on improving accuracy for each: (1) Masc/Fem NP Agr and (2) Masc/Fem Nouns. Therefore, when error tabulations were made for noun phrases like the one seen in number 1, one error was counted for Masc/Fem nouns because the noun was not written in the correct gender. However, no errors were counted for Masc/Fem NP Agr because all the linguistic material within the noun phrase in number 1 agrees in gender, even though the wrong gender is used across all the linguistic materials inside this noun phrase.

Creating the new category of Masc/Fem Nouns required that we categorize it as either a binary or non-binary option construction. When participants encountered the mf symbol, they were required to make a binary choice when rewriting the Spanish nouns in the final draft, and the WCF symbol *mf* provided does, in effect, provide the correct answer with no other options for correct usage. That is, if a student receives an *mf* symbol above the word *plumo*, he or she should know to change it to *pluma*. There are no other options for correct, regular usage of masculine and feminine in Spanish. According to the binary vs. n-ary (n > 2) hypothesis then, we should have seen improvement to participants' selection of the correct masculine/feminine noun endings across elicitation tasks as well as agreement across noun phrases. We did not, and this

finding led to the stipulation of "non-idiosyncratic" being added to the binary option stipulation.

Yet another tabulation decision had to be made regarding disagreeing gender morphology within a noun phrase, as seen in number 2:

2. *la plum**o** roj**a**
   the pen   red
   the red pen

Here, an argument could be made that there are two Masc/Fem NP Agr errors within the noun phrase. Instead of calculating two errors here, however, a decision was made to simply count agreement across noun phrases as either correct or incorrect, so number 2 would be counted as one error in noun phrase agreement, and number 1 would not be counted as an error in noun phrase agreement. This decision was made because the *mf* symbol was initially incorporated into the study to examine if it would help students improve agreement across noun phrases, and it was the underlying intuition of both the researcher and the L2-Spanish teacher who helped mark the writing samples with WCF that this was the duty of the *mf* symbol.

Having said all that, a simple count of *all* masculine/feminine errors was also tabulated across all four elicitation tasks to determine if providing the *mf* symbol helped in the reduction of overall masculine/feminine errors for experimental groups compared to control groups. Consequently, the results in Chapter 7 include the following analyses of masculine/feminine errors:

- overall occurrences of *mf* errors/symbols across the elicitation tasks
- *mf* errors/symbols' effect on correct noun morphology

- *mf* errors/symbols' effect on agreement of all linguistic materials within the noun phrase

Of course, there were more *overall* masculine/feminine errors than noun morphology errors because there were more articles and adjectives modifying nouns than the nouns themselves. On the other hand, overall masculine/feminine errors and masculine/feminine noun phrase agreement errors were about the same count because where there was even one error in agreement in the noun phrase, it made the entire noun phrase incorrect. However, examples such as number 1 did show up in the data. That is, some noun phrases were completely incorrect with regards to the gender chosen, but all the linguistic materials within the noun phrase did match as the same gender.

Once the masculine/feminine data was teased apart to account for the different types of knowledge actually under examination, the researcher realized that the same could be done for Spanish singular/plural errors. Therefore, singular/plural data was coded (1) Sing/Pl and (2) Sing/Pl NP Agr where Sing/Pl simply counts *all* occurrences of singular/plural errors and Sing/Pl NP Agr counts noun phrases with any number disagreement as an error and any noun phrase with no disagreement in number within the noun phrase as correct. Consequently, the results in Chapter 7 include the following analyses of singular/plural errors:

- overall occurrences of *num* errors/symbols across the elicitation tasks
- *num* errors/symbols' effect on agreement of all linguistic materials within the noun phrase

These two data sets were similar as well. However, there were examples of participants

making errors within a noun phrase but all materials agreeing in number across the noun

phrase, as seen in number 3:

<div align="center">

3. **la**   plum**a**
     the pen
     the pens

</div>

In cases like these, the correct usage was *las plumas* because the participant should have

used the plural of both the Spanish word for pen and the article preceding this noun. This

was tabulated as two overall singular/plural errors, but the noun phrase was not tabulated

as an error because it agrees in number. Again, such tabulations of singular/plural errors

allow us to examine proceduralization of two different types of knowledge: overall

singular/plural knowledge and singular/plural agreement across noun phrases.

Finally, this leads to one more point regarding the Spanish noun phrase and the

way errors were marked with WCF within the noun phrases. The noun phrases in the

participant writing samples were marked heavily with a variety of symbols. Many noun

phrases were packed with red ink symbols, as seen in number 3:

<div align="center">

art
mf   *mf*
num   *mf*   *num*
3. el   plumos   roja

</div>

Notice here that "el" is also marked with the *art* symbol to indicate an article error. The

"el" here should be the Spanish article "las" (feminine and plural) to match the correct

production of plumas (corrected as feminine and plural). Marking the error with *mf, num*

and *art* appears to have convoluted the data by counting the same error twice or three

times. That is, "el" is the wrong article, but it is wrong because it is it is the wrong

<div align="center">105</div>

number and the wrong gender, errors already pointed out by the symbols *num* and *mf* respectively. This issue will be addressed in detail in the results in Chapter 7.

A one-way repeated measures of analyses of variance (ANOVA) was chosen to examine the error rate changes/differences in error reduction tendencies (the dependent variable) across the four elicitation tasks between control groups and experimental groups. One-way repeated measures ANOVA is a statistical technique that is useful in investigating the changes in mean scores across three or more time points for correlated samples (Gravetter & Wallnau, 2004). The following comparisons were evaluated in order to find answers to the four research questions above:

1. Experimentals versus Controls— Overall Error
2. Experimental versus Controls— *A Comparison of Category Types (i.e., Linguistic, Word Usage, Mechanical)*
3. Experimentals versus Controls—Individual *Linguistic* Error Types

Errors were tallied from within the first 150 words of each participant's paragraphs, since every participant wrote at least 150 words on each elicitation task (some did not attain the 200 word requirement).

Descriptive statistics and statistical analysis for the data can be seen in Chapter 7. Statistical analysis of the accuracy increases/error reductions between experimental and control groups for total errors, error category types, and individual linguistic errors across the four elicitation tasks uses a one-way repeated ANOVA analysis and the following statistical hypotheses and rules:

> **Null Hypothesis**: There is no difference between experimental and control groups abilities to reduce errors over time for total errors, error category types, and/or individual linguistic errors. In other words, experimental and control

groups behave similarly in their error reduction tendencies over time for total errors, error category types, and/or individual linguistic errors.

Alternative Hypothesis: There is a significant difference between experimental and control groups abilities to reduce errors over time for total errors, error category types, and/or individual linguistic errors. In other words, experimental and control groups behave differently in their error reduction tendencies over time for total errors, error category types, and/or individual linguistic errors.

Rules for Rejecting the Null Hypothesis: If F value is less than 0.05 (95% confidence level of the results), then the difference is significant; otherwise, accept the null hypothesis.

This tabulation and analysis of errors allow us to determine which symbols are effective in transitioning L2 students from declarative knowledge of grammatical constructions to proceduralized knowledge of said constructions via correction practice. According to the binary vs. n-ary (n > 2) hypothesis, Spanish singular/plural and masculine/feminine agreement WCF symbols (*num* and *mf* respectively) should have provided the necessary declarative knowledge to ensure students make the correct revisions to these binary pieces of grammar during practice, and this should lead to proceduralization of the rules for these constructions as demonstrated by experimental participants' significant accuracy gains relative to their control counterparts. This should be the case for singular/plural overall errors and singular/plural agreement across noun phrases. Likewise, masculine/feminine overall errors, noun morphology errors, and agreement across noun phrases should also reduce at a significant rate for experimental groups compared to control groups. On the other hand, Experimental groups' Spanish subject/verb agreement, verb tense, and article usage should not improve significantly better than control groups because of their increased complexity and WCF's inability to elucidate these underlying complexities.

Findings for the relative performance of the experimental and control groups are presented in the next chapter both in numerical and graphical formats. For the purposes of this study, statistically significant accuracy increases over the four elicitation tasks for experimental group participants relative to control group participants represent proceduralized knowledge of individual constructions. Because the use of WCF was the only difference between the experimental (received WCF) and the control groups (did not receive WCF), differences in participants' accuracy increases can be attributed to WCF. Graphic representations of error reduction tendencies that include flat lines or lines that slope upward across the four elicitation tasks imply a lack of proceduralization. Lines that slope downward represent strong reductions in errors and proceduralization.

.

CHAPTER SEVEN: L2- SPANISH WCF STUDY RESULTS

Data comparing the experimental and control groups' reduction rates generally support the alternative statistical hypothesis and the current study's binary vs. n-ary (n > 2) research hypothesis, with a few stipulations. Descriptive and statistical statistics are presented together for each category of error analysis.

Descriptive Statistics and Statistical Analysis of Findings

We begin with a macro examination of the data in order to determine WCF's overall effectiveness for total errors and for error categories (i.e., linguistic errors, mechanical errors, and word usage errors). As discussed in the Chapter 1 introduction and the Chapter 2 literature review, this is where many WCF researchers end their investigation of WCF. They examine the data and make the claim that WCF is either effective or ineffective. Such investigations conceal how WCF actually works to help L2 students to become more accurate writers. It also obscures precisely for which grammatical categories WCF can be an effective tool for improving grammatical accuracy. Because the vast majority of WCF studies provide no explanation for *how* or *why* WCF works or does not work, debates over WCF effectiveness have continued for decades with no resolution. Some say, "It works, see?" Others say, "It doesn't work, see?"

Consequently, after the macro examination of the data is concluded, data from individual grammatical categories is examined in order to analyze the effectiveness of WCF on these individual constructions. It is this microanalysis of the data that helps to answer the research questions and to confirm or refute the research hypothesis for the current study. Still, beginning the analysis at the macro level and only then probing more deeply into the data in a step-by-step approach is instructive. This type of analysis demonstrates how confusion about WCF effectiveness and the implementation of WCF occurs when only macro data is analyzed, as it is so often in the literature. Indeed, when more detailed analyses are conducted, it becomes clear that WCF should only be used on constructions for which it will work, the simple constructions.

A comparison of overall error reduction rates across the four elicitation tasks between experimental and control groups demonstrates a stark contrast between the participant error reduction tendencies in the two groups. These results do not help to answer the specific research questions for this study, nor do they confirm or refute the binary vs. n-ary (n > 2) hypothesis. However, it is important to at least start with an examination of total error result in order to confirm prior research on WCF overall effectiveness and to demonstrate that examinations of overall effectiveness conceal important findings. Then, we can probe into the data in order to examine WCF effectiveness more thoroughly and according to the research questions and hypothesis herein.

Total Errors. Results show a statistically significant difference in the reduction of total errors over the four elicitation tasks between participants from the experimental

group and participants from the control group (F = 5.405, p = 0.026) with experimental groups increasing overall accuracy significantly better than control groups over the 12-week elicitation period (with a statistical significance level of < .05). Figure 7.1 presents descriptive statistics, showing changes in the total number of errors across the four elicitation tasks. Table 7.1 also presents descriptive statistics, showing the mean error rates by experimental and control groups across the four elicitation tasks as well as the standard deviation from the mean for each elicitation task. Figure 7.2 provides a graphic representation of the data from Table 7.1.

All results will be provided in this Figure/Table/Figure format. Presenting the data in this way provides the opportunity to compare *numerical* changes to the amount of errors (e.g. 654 to 506 between tasks 1 and 2) across the elicitation tasks for experimental and control groups. It also allows us to examine more closely the changes to the means of errors for experimental and control groups both in a numerical format and a visual representation. Moreover, the one-way repeated ANOVA P Values are based on a comparison of differences to the changes of the means between the experimental and control group participants. Therefore, this style of data presentation allows for a full picture of the findings the data represents.
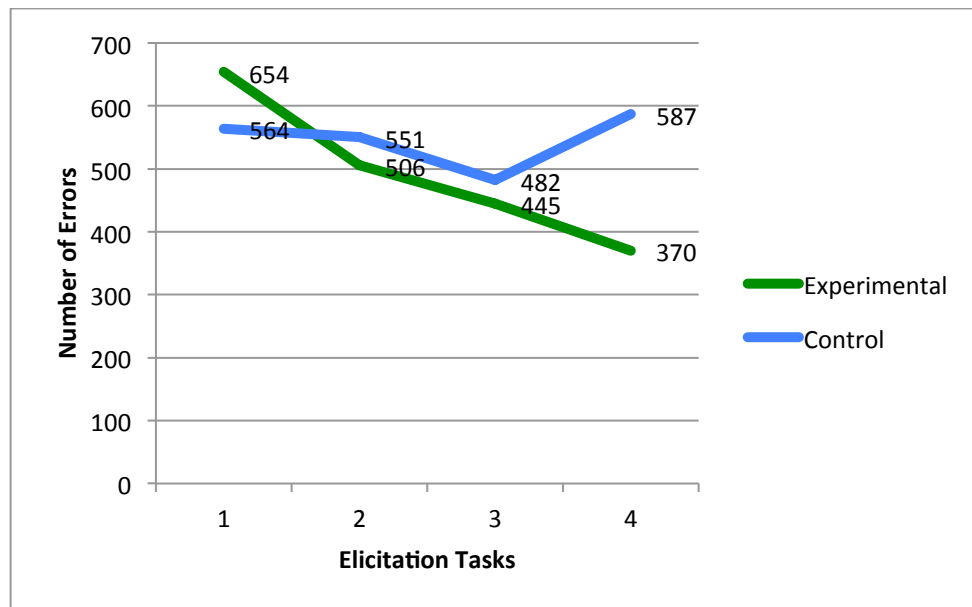
Figure 7.1: Total Errors (Includes all Error Types)
Descriptive statistics for numerical changes across elicitation tasks

In Figure 7.1, we see the four elicitation tasks on the x-axis and the number of total errors represented on the y-axis. It becomes immediately clear that the green line representing the experimental group's changes to the number of errors they commit over time slopes down more sharply than the blue line representing the control group's changes to the number of errors they commit over time. Generally speaking, strong downward sloping lines throughout the data sets represent strong error reduction tendencies across the elicitation tasks, and this, in turn, represents proceduralization of knowledge according to how we are defining/operationalizing proceduralization in this study (i.e., accuracy increases/error decreases). Where lines are flat, slope downward only marginally, or even slop upward; they represent a lack of proceduralization of

112

knowledge over time as demonstrated by a lack of error reduction (or lack of accuracy increases) over time.   Of course, it is the comparison of the differences of these slopes between experimental and control groups that provides us with a level of significance.

Participants in the experimental group reduce their total errors for all error types from 654 to 506 between tasks 1 and 2 (-23%), from 654 to 445 between tasks 1 and 3 (-32%), and from 654 to 370 between tasks 1 and 4 (-43%). This shows a strong, downward sloping tendency in the number of total errors for the experimental group participants. Experimental participants increase their accuracy between the first and fourth elicitation tasks by making 43% fewer errors, demonstrating improvement to overall knowledge of writing conventions for Spanish. The control participants do not demonstrate this same downward sloping tendency in their error reduction tendencies. They demonstrate a modest downward trend from 564 to 551 between tasks 1 and 2 (-2%) and from 564 to 482 between tasks 1 and 3 (-15%). Then, the number of total errors actually increases between tasks 1 and 4 from 564 to 587 (+4%).

This increase on the 4th elicitation task for the control group participants occurs repeatedly for each error type. It also occurs for experimental group participants for error types that are not amenable to positive change via WCF. This could be a result of a number of factors. It could be that control participants become less careful as time goes on because they know that they will not be held accountable for errors. It could be that experimental groups also become less careful on structures for which they continually receive feedback but no discernable declarative knowledge. Another explanation could be that the cause and effect rhetorical form that participants were required to write for the

4th elicitation task was more difficult for them or affected their error rates in some way. Regardless, WCF on binary, non-idiosyncratic forms does appear to have kept experimental group participants from mirroring control group participants' uptick in errors on the 4th elicitation task.

Table 7.1 compares mean scores for total errors between experimental and control groups across the four elicitation tasks, and Figure 7.2 shows the graphic representation of the values in Table 7.1.

Table 7.1: Total Errors (Includes all Error Types)
Descriptive statistics for mean test scores by group and elicitation task

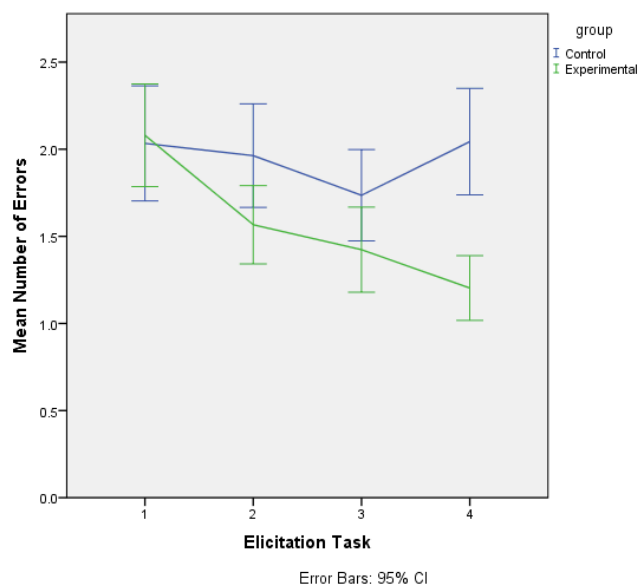| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	2.1	2.6	1.6	2.0	1.4	2.1	1.2	1.6
Control	19	2.0	2.9	2.0	2.6	2.7	2.3	2.0	2.7

Figure 7.2: Effectiveness of WCF on Total Errors over Time

Findings on overall error reductions appear to provide a strong endorsement for the use of indirect WCF in the L2-Spanish classroom. However, this is precisely the type of analysis that has led to the confusion and debates over WCF effectiveness in the literature. Although researchers and L2-writing instructors may look at these results as a confirmation that this L2-teaching strategy is effective and should be implemented, many of the fundamental issues raised by opponents to WCF are not addressed with such a broad analysis. That is precisely why more detailed examinations of the data are necessary for a clearer understanding of WCF utility.

Total Linguistic Errors. Next, we examine WCF effectiveness on total linguistic errors, which includes eight separate grammatical error types: singular/plural

total occurrences, singular/plural noun phrase agreement, masculine/feminine total occurrences, masculine/feminine noun phrase agreement, masculine/feminine noun morphology, subject-verb agreement, verb tense, and articles. These error types are all examined individually below in order to answer this study's research questions and confirm or refute the research hypothesis, but it is imperative that as we dig deeper into the data, we do so one step at a time. Results for total linguistic errors show a statistically significant difference in the reduction of total linguistic errors over the four elicitation tasks between participants from the experimental group and participants from the control group ($F = 5.716$, $p = 0.001$).
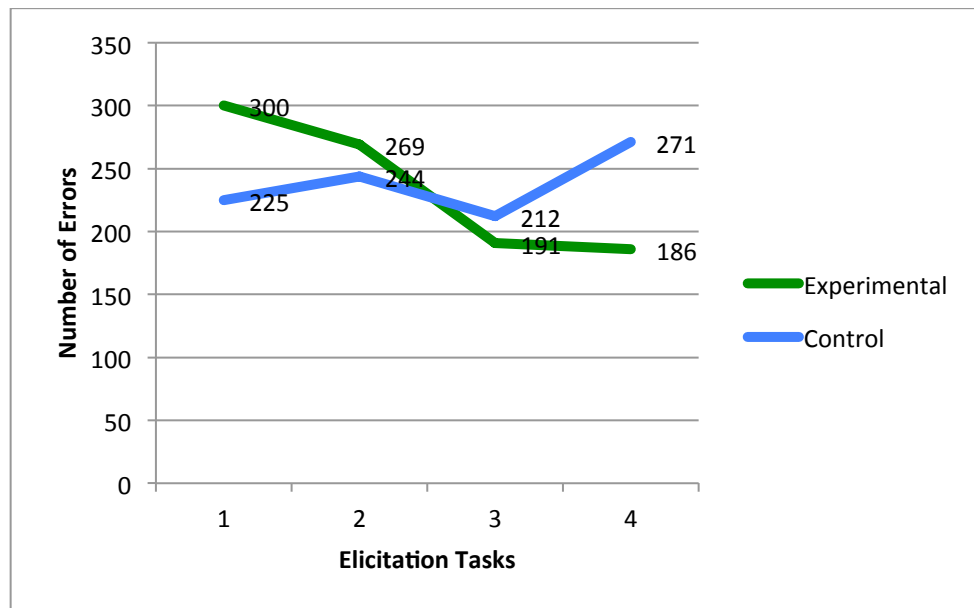

Figure 7.3: Total Linguistic Errors
Descriptive statistics for numerical changes across elicitation tasks

Data in Figure 7.3 demonstrates a distinct difference in participant behaviors between the experimental and control groups. Once again, experimental participants' error reduction tendencies slope sharply down while control groups actually increase over the same elicitation period. Participants in the experimental group reduce their total linguistic errors from 300 to 269 between tasks 1 and 2 (-10%), from 300 to 191 between tasks 1 and 3 (-36%), and from 300 to 186 between tasks 1 and 4 (-38%). Control participants' data demonstrate either increases in errors or only modest downward trends, especially relative to their experimental counterparts. Their linguistic errors increase from 225 to 244 between tasks 1 and 2 (+8%), decrease marginally from 228 to 212 between tasks 1 and 3 (-6%), and increase again from 228 to 271 between tasks 1 and 4 (+20%).

Table 7.2 compares mean scores for total errors between experimental and control groups across the four elicitation tasks, and Figure 7.4 shows the graphic representation of the values in Table 7.2.

Table 7.2: Total Linguistic Errors
Descriptive statistics for mean test scores by group and elicitation task

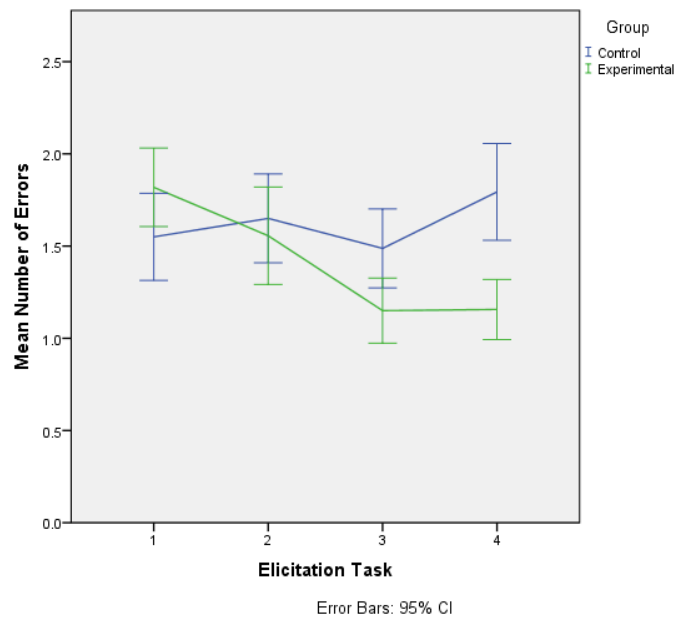| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	1.82	1.36	1.56	1.69	1.15	1.13	1.16	1.04
Control	19	1.55	1.51	1.65	1.54	1.49	1.37	1.79	1.68

Figure 7.4: Effectiveness of WCF on Total Linguistic Errors over Time

Once again, these findings could be seen as an endorsement of indirect WCF for improving L2-students linguistic accuracy. However, we still do not know for which constructions WCF provides sufficient declarative knowledge and for which it does not.

Total Mechanical Errors. Now we will examine WCF effectiveness on total mechanical errors, which include capitalization, run-on sentences, fragments, punctuation, and spelling. Results for total mechanical errors show a statistically significant difference in the reduction of total mechanical errors over the four elicitation tasks between participants from the experimental group and participants from the control group (F = 5.207, p = 0.002).
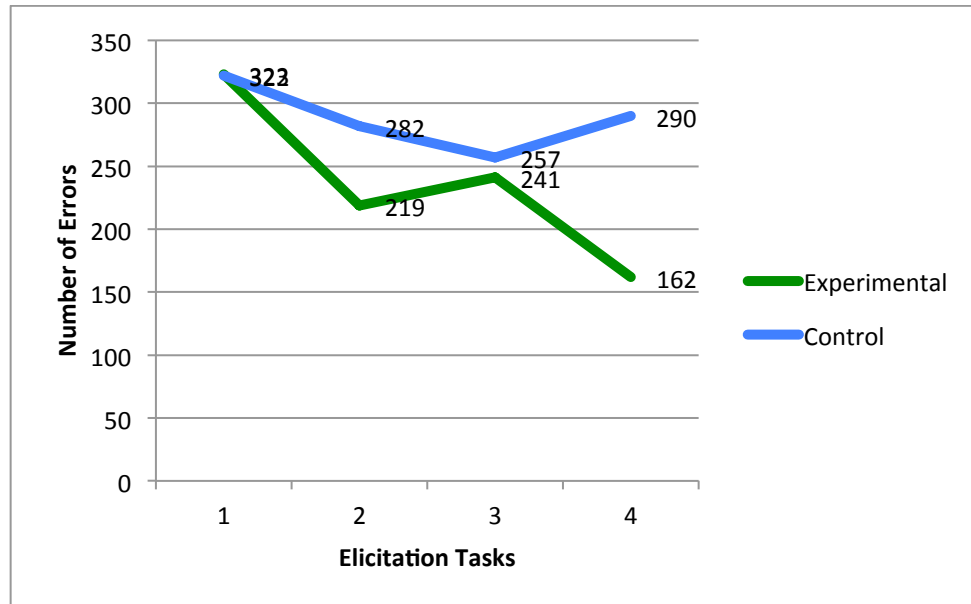
Figure 7.5: Total Mechanical Errors
Descriptive statistics for numerical changes across elicitation tasks

Here again, we see sharper reduction tendencies for experimental groups relative to control groups. Participants in the experimental group reduce their total mechanical errors from 323 to 219 between tasks 1 and 2 (-32%), from 323 to 241 between tasks 1 and 3 (-25%), and from 323 to 162 between tasks 1 and 4 (-50%). Control participants' data demonstrate modest downward trends relative to the experimental groups' data. Their mechanical errors decrease nominally from 322 to 282 between tasks 1 and 2 (-12%), decrease again from 322 to 257 between tasks 1 and 3 (-20%), and decrease slightly from 322 to 290 between tasks 1 and 4 (-10%).

119

Table 7.3 compares mean scores for total errors between experimental and control groups across the four elicitation tasks, and Figure 7.6 shows the graphic representation of the values in Table 7.3.

Table 7.3: Total Mechanical Errors
Descriptive statistics for mean test scores by group and elicitation task

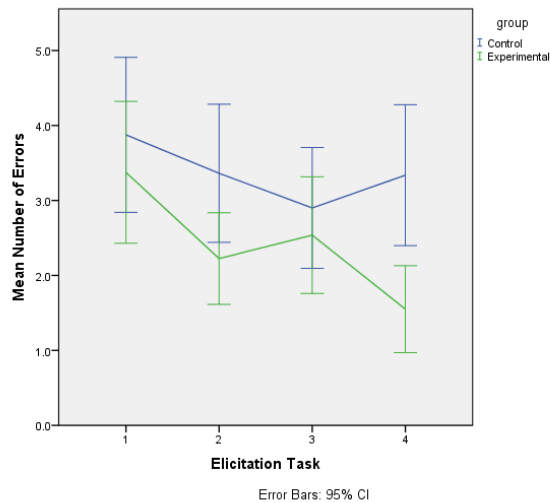| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	3.38	4.25	2.23	2.75	2.54	3.50	1.55	2.61
Control	19	3.92	4.65	3.41	4.15	2.94	3.63	3.38	4.23

Figure 7.6: Effectiveness of WCF on Total Mechanical Errors over Time

Examining this data proves instructive regarding how an analysis of total errors either within a category or overall error tabulations across all error categories can obscure the actual utility of WCF. Remember that the error category "mechanical errors" includes capitalization, punctuation, run-on sentences, fragments, and spelling. One might be inclined to conclude that all these error types reduced at statistically significant levels based on the data above. That, however, is not the case. Within this category, there are very few run-on and fragment errors to begin with. Furthermore, run-on, fragment, and spelling errors generally do not reduce meaningfully for either experimental or control groups (i.e. their reduction tendencies remained flat). Given that capitalization and punctuation represent the majority of the errors within the category and given that experimental group participants reduce these two error types significantly better compared to control group participants, it can be concluded that it is the combination of data from these two categories that demonstrates statistically significant differences between experimental and control groups, not the other three error types. Of the mechanical 945 mechanical errors committed by the experimental group participants, 755 are capitalization errors (292) and punctuation errors (463) or a total of 80% of the total mechanical errors for the group. Of the 1151 mechanical errors committed by the control group participants, 958 are capitalization errors (367) and punctuation errors (591) or a total of 83% of the total mechanical errors for the group. These large percentages for the two error types relative to the other three skew the data. Experimental participants' reductions of these two error types are significantly better relative to the

control group participants' reductions, and data from these two error types is what give us a statistically significant difference between the groups for the whole category.

Upon closer examination of the writing samples from all four elicitation tasks, it becomes clear that the vast majority of improvements to capitalization errors are made by experimental participants simply capitalizing the first words of sentences and that the vast majority of improvements to punctuation errors are made by experimental participants providing ending punctuation at the ends of sentences. This is knowledge these upperclassmen undoubtedly had before receiving WCF. However, it appears that they become conscious of the fact that they would be corrected for such errors, and they begin to provide the correct capitalization and punctuation as a result of WCF. The control group participants, however, do not recognize the importance of accuracy on such errors and continue to not provide correct capitalization and punctuation.

Total Word Usage Errors. Now we will examine WCF effectiveness on total word usage errors, which include wrong word errors and word form errors. Results for total word usage errors do not show a statistically significant difference in the reduction of total word usage errors over the four elicitation tasks between participants from the experimental group and participants from the control group ($F = 0.882$, $p = 0.450$). This is probably a result of having so few occurrences of both types of errors across the data set. Still, the descriptive statistics at least provide some indication that the experimental group outperforms the control group even if statistical significance was not attained.
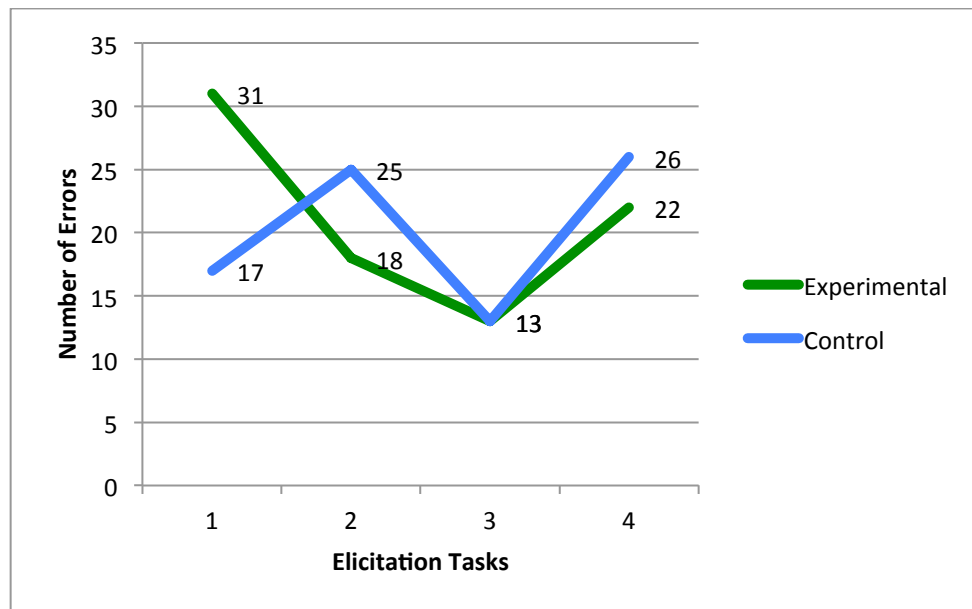
Figure 7.7: Total Word-Usage Errors
Descriptive statistics for numerical changes across elicitation tasks

Participants in the experimental group reduce their total word usage errors from 31 to 18 between tasks 1 and 2 (-42%), from 31 to 13 between tasks 1 and 3 (-58%), and from 31 to 22 between tasks 1 and 4 (-29%). It could also be the increase in errors at elicitation task 4 relative to the downward trend seen for tasks 2 and 3 produces the result as not significantly different between both groups. Control participants increase their total word usage errors from 17 to 25 between tasks 1 and 2 (+47%), reduce errors from 17 to 13 between tasks 1 and 3 (-24%), and again increase errors from 17 to 26 between tasks 1 to 4 (+53%).

Table 7.4 compares mean scores for total errors between experimental and control groups across the four elicitation tasks, and Figure 7.8 shows the graphic representation of the values in Table 7.4.

Table 7.4: Total Word-Usage Errors
Descriptive statistics for mean test scores by group and elicitation task

Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	1.05	1.14	0.72	0.87	0.67	1.05	0.87	1.13
Control	19	0.83	0.98	0.90	1.00	0.82	1.14	0.95	1.17
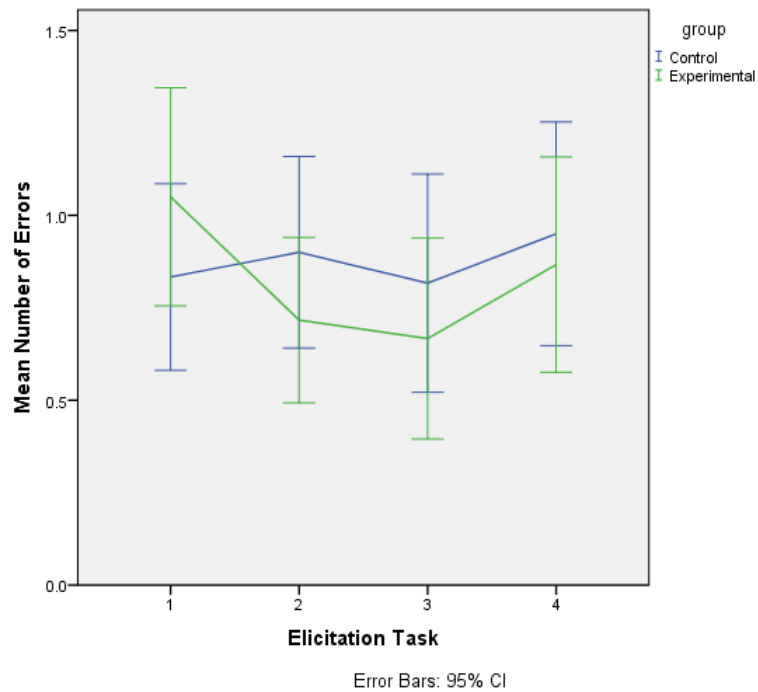
Figure 7.8: Effectiveness of WCF on Total Word-Usage Errors over Time

Findings from this macro examination of the data generally confirm empirical findings from WCF studies in the academic literature and from Wagner (2013) with regards to WCF's overall effectiveness. However, this examination of WCF provides no explanation for how WCF actually works. It also obscures findings that point to reasons for its variable effectiveness on different grammatical forms. The analysis of the individual grammatical categories below demonstrates the importance of examining WCF under SAT, which gauges its effectiveness by its ability to impart declarative knowledge of individual constructions.

Now we will examine the data in more detail by analyzing WCF effectiveness for each grammatical construction. It is this data that helps to answer the current study's research questions and to either confirm or refute binary vs. n-ary (n > 2) research hypothesis. There are a total of eight individual grammatical categories under analysis: singular/plural total errors, singular/plural noun phrase agreement errors, masculine/feminine total errors, masculine/feminine noun phrase agreement errors, masculine/feminine noun morphology errors, subject-verb agreement errors, verb tense errors, and articles errors. According to the binary vs. n-ary (n > 2) hypothesis, it was predicted that singular/plural total errors, singular/plural noun phrase agreement, masculine/feminine total errors, masculine/feminine noun phrase agreement errors, and masculine/feminine noun morphology errors would be amenable to positive change via the use of WCF. That is, experimental groups would significantly outperform control groups in their abilities to reduce these error types over the four elicitation tasks (i.e., become more accurate and demonstrate proceduralization of the forms). On the other hand, Spanish subject-verb agreement, verb tense, and article usage were predicted not to be amenable to change via the use of WCF because these constructions in Spanish require more than a binary choice for proper error correction and/or usage.

Overall Singular/Plural Errors. We begin with an analysis of singular/plural total errors. This is a count of all occurrences of singular plural errors or all occurrences of the symbol *num* across the four elicitation tasks. Results for singular/plural total errors do not show a statistically significant difference in the reduction of total singular/plural

errors across the four elicitation tasks between participants from the experimental group and participants from the control group (F = 1.272, p = 0.102).
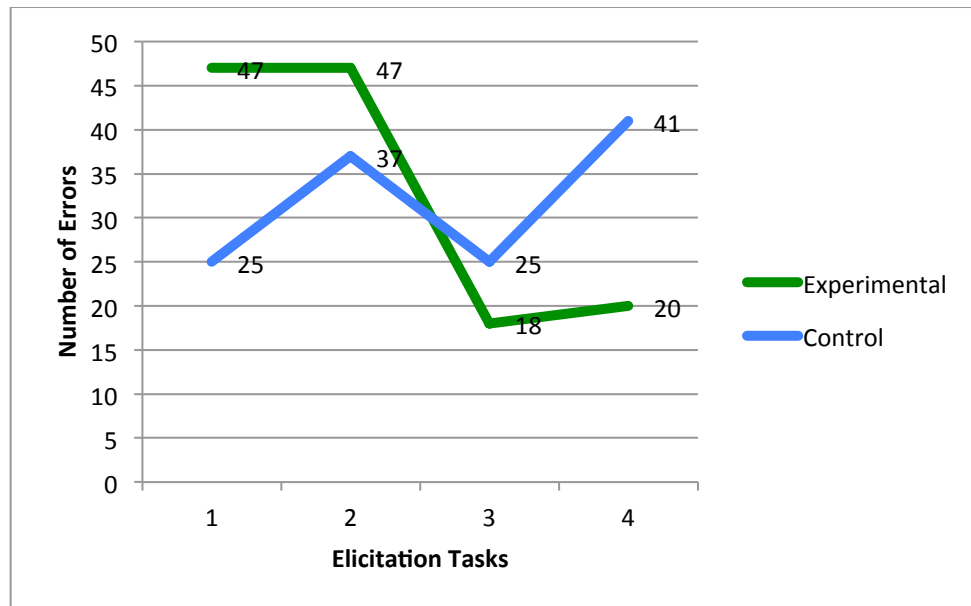
Figure 7.9: Singular/Plural Total Errors
Descriptive statistics for numerical changes across elicitation tasks

Participants in the experimental group actually do not reduce their singular/plural errors between tasks 1 and 2 (0%). They commit exactly 47 errors on both the first and second elicitation tasks. However, they reduce from 47 to 18 between tasks 1 and 3 (-63%), and from 47 to 20 between tasks 1 and 4 (-57%). Control participants increase their total singular plural errors from 25 to 37 between tasks 1 and 2 (+48%) and make no increases or decreases between tasks 1 and 3 (0%). They increase singular plural errors from 25 to 41 between tasks 1 to 4 (+64%).

Table 7.5 compares mean scores for total errors between experimental and control groups across the four elicitation tasks for total singular/plural errors, and Figure 7.10 shows the graphic representation of the values in Table 7.5.

Table 7.5: Singular/Plural Total Errors
Descriptive statistics for mean test scores by group and elicitation task

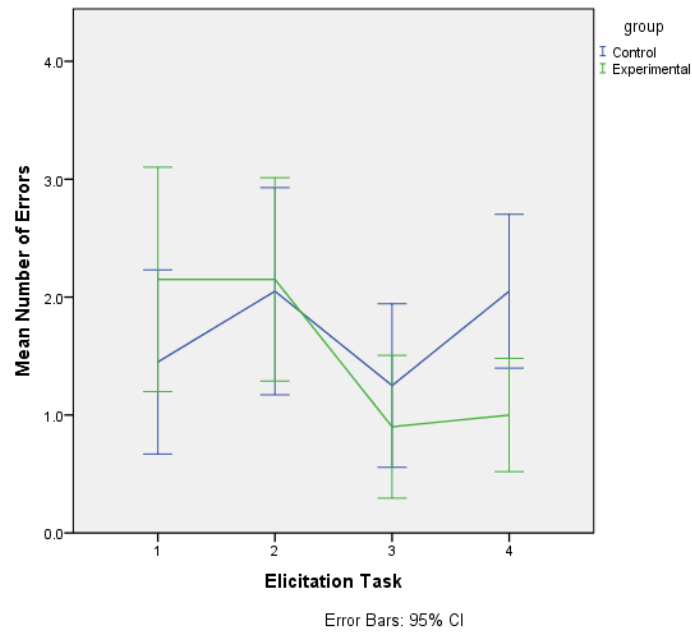| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	2.15	2.03	2.15	1.84	0.90	1.29	1.00	1.03
Control	19	1.45	1.67	2.05	1.88	1.25	1.48	2.05	1.39

Figure 7.10: Effectiveness of WCF on Singular/Plural Total Errors over Time

128

This first analysis of WCF effectiveness on an individual grammatical category does not demonstrate support for the alternative statistical hypothesis or this study's research hypothesis. The binary vs. n-ary (n > 2) predicts that experimental group participants should outperform control group participants on this binary option construction. Descriptive statistics appear to demonstrate that the two groups have different error reduction tendencies across the elicitation tasks, but these differences do not reach a level of statistical significance. This brings us to the research questions and research hypothesis.

> **RQ 1:** Does WCF and the correction process aid in the transition from declarative to proceduralized knowledge of specific grammatical constructions in the writing modality? And if so, for which constructions is this transition observed?

No strong claim can be made that WCF facilitates the transition from declarative to proceduralized knowledge for grammatical category singular/plural.

The analysis of singular/plural overall errors appears to demonstrate that WCF may even be ineffective for some binary option constructions. This would mean that the grammatical complexity threshold addressed by Research Question 2 could be less than two, rendering WCF virtually useless:

> **RQ 2**: Given the answers to RQ1, is there a grammatical complexity threshold beyond which WCF will no longer be operative in supplying the declarative knowledge necessary for L2-writing students to practice and eventually proceduralize?

This binary option construction was not amenable to change via the use of WCF. However, other binary option constructions do appear to be amenable to positive change with the use of WCF.

      Singular/Plural Noun Phrase Agreement Errors. Now, we can examine the effectiveness of the symbol *num* on improving number agreement across the noun phrase. Instead of counting every occurrence of the error/symbol *num*, the number of noun phrases that contain any type of disagreement (i.e., articles, nouns, or adjectives) are tabulated as an error. This tabulation allows us to examine WCF effectiveness for providing declarative knowledge on the relationship among linguistic materials within noun phrases as opposed to only examining the knowledge that "s" must be added to plural nouns. Results for singular/plural noun phrase agreement errors show a statistically significant difference in the reduction of this error type over the four elicitation tasks between participants from the experimental group and participants from the control group (F= 3.25, p = 0.024).
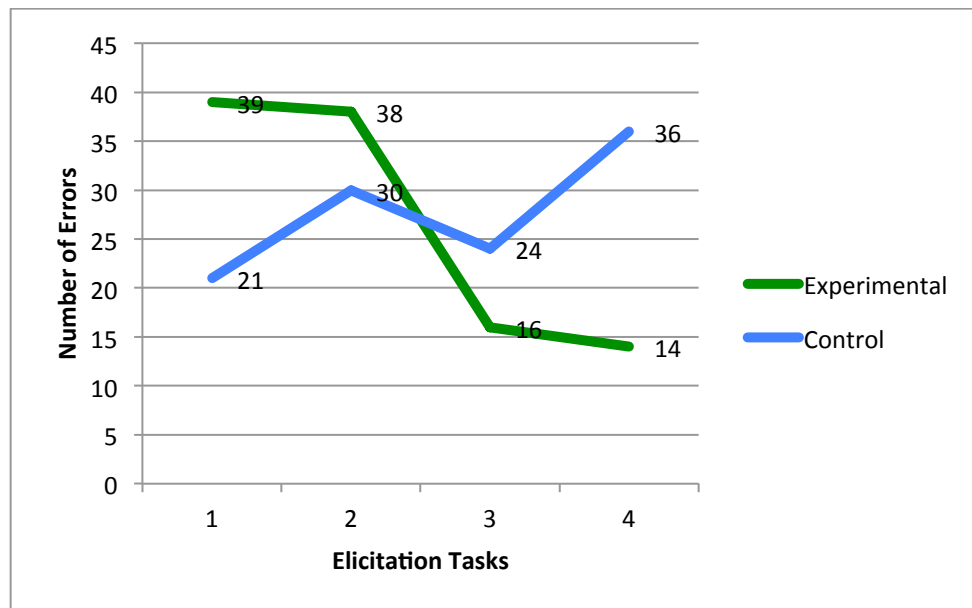
Figure 7.11: Singular/Plural NP Agreement Errors
Descriptive statistics for numerical changes across elicitation tasks

Participants in the experimental group reduce their singular/plural noun phrase agreement errors only marginally from 39 to 38 between tasks 1 and 2 (-3%). However, they reduce from 39 to 16 between tasks 1 and 3 (-59%), and from 47 to 14 between tasks 1 and 4 (-64%). Control participants increase their total singular plural errors from 21 to 30 between tasks 1 and 2 (+43%), from 21 to 24 between tasks 1 and 3 (+14%), and from 21 to 36 between tasks 1 and 4 (+71%).

Table 7.6 compares mean scores for total errors between experimental and control groups across the four elicitation tasks for singular/plural noun phrase agreement, and Figure 7.12 shows the graphic representation of the values in Table 7.6.

Table 7.6: Singular/Plural NP Agreement Errors
Descriptive statistics for mean test scores by group and elicitation task

Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	1.80	1.58	1.85	1.57	0.80	1.06	0.70	0.80
Control	19	1.20	1.24	1.55	1.54	1.20	1.44	1.80	1.24

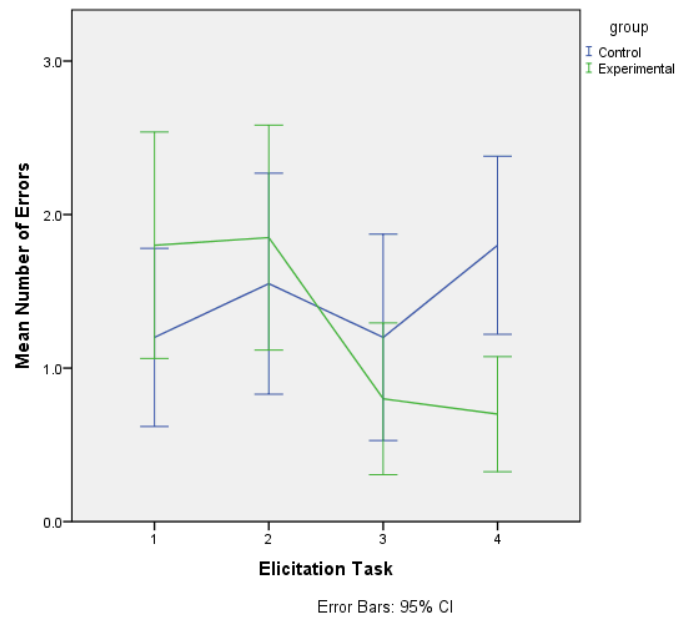***Figure 7.12: Effectiveness of WCF on Singular/Plural NP Agreement Errors over Time***

Here, the answer to research question 1 can be answered in the affirmative. The repetition of the WCF symbol *num* across noun phrases does provide enough declarative knowledge to L2-Spanish students to practice meaningfully, and correction practice leads to proceduralization of the knowledge as demonstrated by experimental groups' ability to

significantly reduce singular/plural noun phrase agreement errors relative to control groups. These findings support the binary vs. n-ary (n > 2) hypothesis. This binary option construction is amenable to positive change via the use of the indirect WCF symbol *num*.

Overall Masculine/Feminine Errors. Now we can take a detailed look at Spanish masculine/feminine agreement. Remember that this data was split into three separate categories for a more detailed examination of how well the WCF symbol mf imparts declarative knowledge of (1) overall masculine usage, (2) masculine/feminine agreement across the noun phrase, and (3) masculine/feminine morphology on the nouns alone. We begin by looking at overall masculine/feminine errors. This is a tabulation of all occurrences of masculine/feminine errors or all occurrences of the symbol *mf* across the four elicitation tasks. Results for masculine/feminine total errors show a statistically significant difference in the reduction of total masculine/feminine errors over the four elicitation tasks between participants from the experimental group and participants from the control group (F = 2.849, p = 0.039).
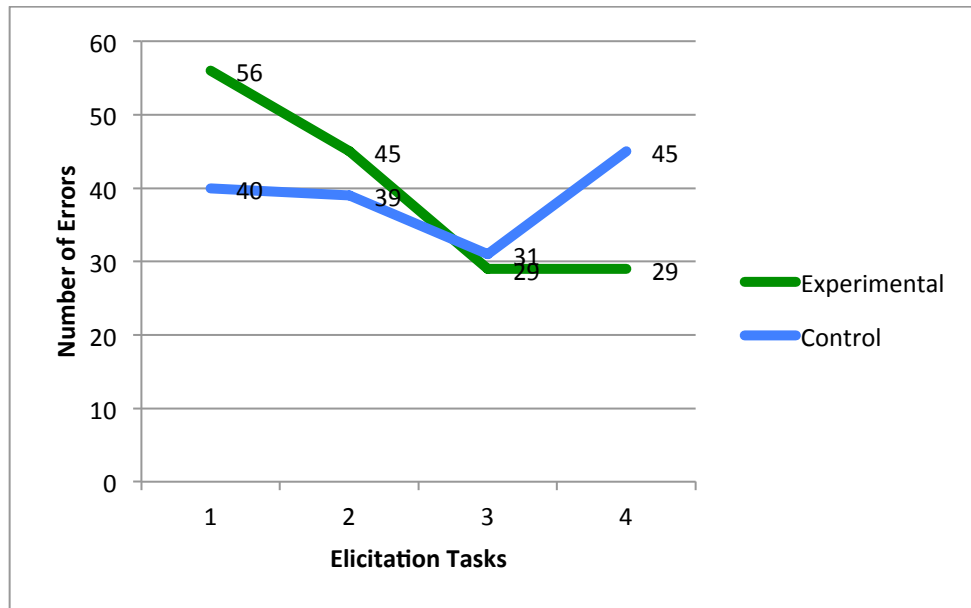
Figure 7.13: Total Masculine/Feminine Errors
Descriptive statistics for numerical changes across elicitation tasks

Participants in the experimental group reduce their total masculine/feminine errors from 56 to 45 between tasks 1 and 2 (-20%). They continue to demonstrate a strong downward reduction tendency, reducing errors from 56 to 29 between tasks 1 and 3 and between tasks 1 and 4 (-48%). Control participants decrease their total masculine/feminine errors from 40 to 39 between tasks 1 and 2 (-2%) and from 40 to 31 between tasks 1 and 3 (-22%). However, they increase their total masculine/feminine errors from 40 to 45 between tasks 1 to 4 (+13%).

Table 7.7 compares mean scores for total errors between experimental and control groups across the four elicitation tasks for total masculine/feminine errors, and Figure 7.14 shows the graphic representation of the values in Table 7.7.

Table 7.7: Total Masculine/Feminine Errors
Descriptive statistics for mean test scores by group and elicitation task

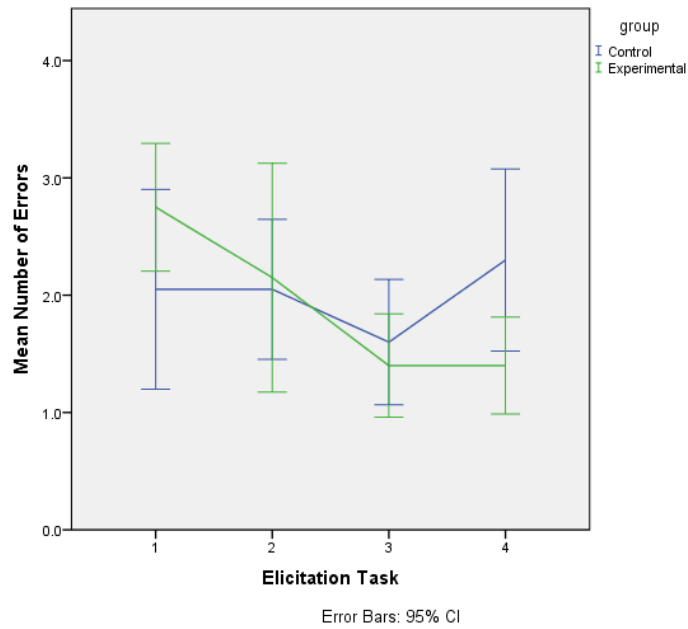| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	2.75	1.16	2.15	2.08	1.40	0.94	1.40	0.88
Control	19	2.05	1.82	2.05	1.28	1.60	1.14	2.30	1.66

Figure 7.14: Effectiveness of WCF on Total Masculine/Feminine Errors over Time

The WCF *mf* symbol appears to provide sufficient declarative knowledge of overall masculine/feminine usage. Practicing corrections to this error type also appears to facilitate a transition from declarative to proceduralized knowledge as demonstrated by

experimental group participants' significant error reductions relative to control group participants' reductions. These findings support the binary vs. n-ary (n > 2) hypothesis.

Masculine/Feminine Noun Phrase Agreement Errors. Now for an examination of how the repetition of *mf* symbol across noun phrase affects participants' reduction rates. Here, if any linguistic materials within the noun phrase did not agree, an error was tabulated. Where all linguistic materials agreed, no error was tabulated, even if the selection of masculine or feminine was incorrect. Results for masculine/feminine noun phrase agreement errors show a statistically significant difference in the reduction of agreement errors over the four elicitation tasks between participants from the experimental group and participants from the control group ($F = 4.052$, $p = 0.008$).
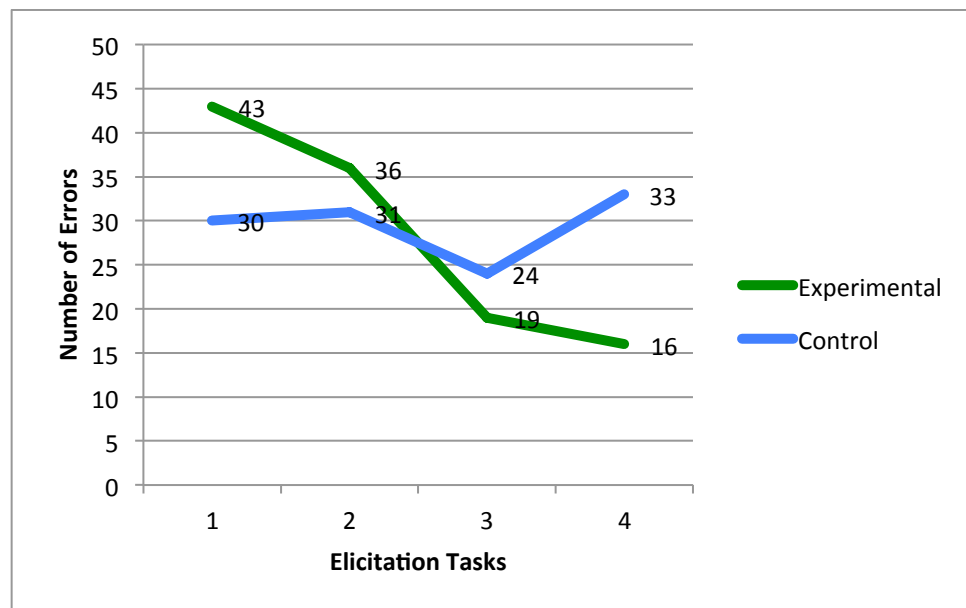
Figure 7.15: Masculine/Feminine NP Agreement Errors
Descriptive statistics for numerical changes across elicitation tasks

Participants in the experimental group reduce their masculine/feminine noun phrase agreement errors from 43 to 36 between tasks 1 and 2 (-17%). They reduce errors from 43 to 19 between tasks 1 and 3 (-56%) and from 43 to 16 between tasks 1 and 4 (-63%). Control participants increase their total masculine/feminine noun phrase agreement errors from 30 to 31 between tasks 1 and 2 (+3%), decreases errors marginally from 30 to 24 between tasks 1 and 3 (-20%), and again increase errors from 30 to 33 between tasks 1 to 4 (+10%).

Table 7.8 compares mean scores for total errors between experimental and control groups across the four elicitation tasks for masculine/feminine noun phrase agreement, and Figure 7.16 shows the graphic representation of the values in Table 7.8.

Table 7.8: Masculine/Feminine NP Agreement Errors
Descriptive statistics for mean test scores by group and elicitation task

Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	2.10	0.91	1.70	1.34	0.90	0.72	0.75	0.55
Control	19	1.55	1.19	1.65	1.09	1.25	0.91	1.70	1.03
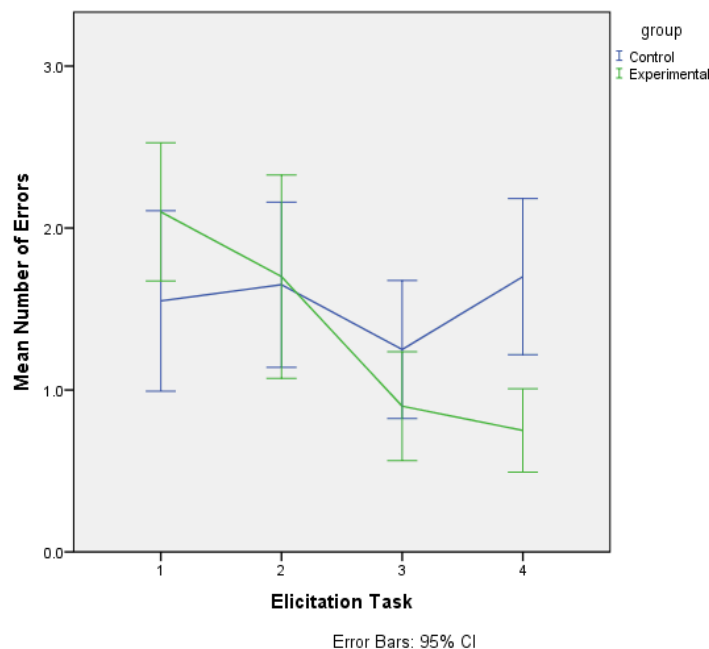
Figure 7.16: Effectiveness of WCF on Masculine/Feminine NP Agreement Errors over Time

Once again, the WCF *mf* symbol appears to provide sufficient declarative knowledge for meaningful practice to begin. Moreover, the repetition of the symbol across noun phrase errors appears to have provided sufficient knowledge that linguistic materials across the noun phrase must agree. This symbol is effective, and the grammatical category it represents can be improved via the use of WCF and the correction practice procedure. This binary option construction is amenable to positive change.

Masculine/Feminine Noun Morphology Errors. Now we can examine data on the effectiveness of the indirect WCF *mf* symbol on helping L2-Spanish students acquire

the correct noun morphology for correct masculine/feminine usage of individual nouns. According to the binary vs. n-ary (n > 2) hypothesis, this grammatical construction should be amenable to change with the use of WCF. It is a binary option construction. Nouns are masculine or feminine, and nouns are generally either marked with an "o" ending for masculine or an "a" ending for feminine. If an ending is written in error by an L2-Spanish student, the "other option" is the correct answer, and crucially there are no other options for correct usage with some minor exceptions.

There are also nouns that end in consonants that require only the consonant ending for masculine and an addition of "a" for the feminine ending (e.g., el profesor and la profesora). Also, nouns that end in "cion" are generally feminine, but the only evidence of its gender is in the preceding article "la" (e.g., la habitacion- the room). If for example a participant wrote "el habitacion," they received an *art* symbol and a *mf* symbol over the article "el" but no WCF over the word "habitacion." This type of error was tabulated as one overall masculine/feminine error (see above) and one masculine/feminine noun phrase agreement error (see above). However, it was not calculated as a noun morphology error.

In this particular examination, errors in noun morphology representing participants' lack of knowledge of the gender of particular nouns is examined. For example, a participant wrote "la dormitoria," but this word is a masculine word in Spanish and should be written "el dormitorio." This represents a lack of knowledge of the gender for this word. The participant received two *mf* symbols here, which were tabulated as two overall errors (see above). Zero errors were tabulated for agreement

because the linguistic materials in the noun phrase (i.e., the article "la" and the ending "a" at the end of "dormitoria") agree in gender even if the gender selected is the wrong gender. Most important for the following analysis, one error was tabulated for the wrong morphology on the noun (i.e., "a" instead of "o").

This analysis allows us to determine if the *mf* symbol helps develop participants' knowledge of different nouns' genders. According to the data, the *mf* symbol is not effective in providing sufficient knowledge of the gender of Spanish words. Results for masculine/feminine noun errors do not show a statistically significant difference in the reduction of specific noun gender errors over the four elicitation tasks between participants from the experimental group and participants from the control group ($F = 0.368$, $p = 0.776$). This finding contradicts the binary vs. n-ary (n > 2) hypothesis. Errors for this binary option construction did not reduce differently between the experimental and control groups.
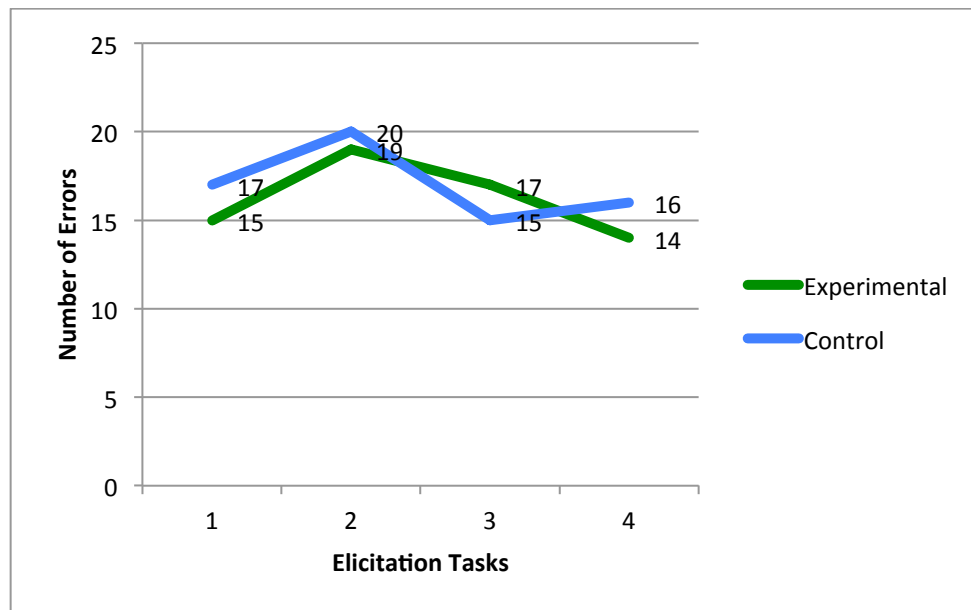
Figure 7.17: Masculine/Feminine Noun Morphology Errors
Descriptive statistics for numerical changes across elicitation tasks

Looking at the line Figure 7.17, it becomes immediately clear that experimental and control groups behave similarly in their reduction of masculine/feminine morphological gender errors as applied to the noun alone. Neither group makes noteworthy reductions to this error type across the four elicitation tasks. Participants in the experimental group actually increase their masculine/feminine noun errors from 15 to 19 between tasks 1 and 2 (+26%) and from 15 to 17 between tasks 1 and 3 (+13%). They reduce errors only marginally from 15 to 14 between tasks 1 and 4 (-7%). Control participants increase their total masculine/feminine noun errors from 17 to 20 between tasks 1 and 2 (+17%) and from decrease errors only nominally from 17 to 15 between tasks 1 and 3 (-12%) and from 17 to 16 between tasks 1 to 4 (-6%).

Table 7.9 compares mean scores for total errors between experimental and control groups across the four elicitation tasks for masculine/feminine noun morphology, and Figure 7.18 shows the graphic representation of the values in Table 7.9.

Table 7.9: Masculine/Feminine Noun Morphology Errors
Descriptive statistics for mean test scores by group and elicitation task

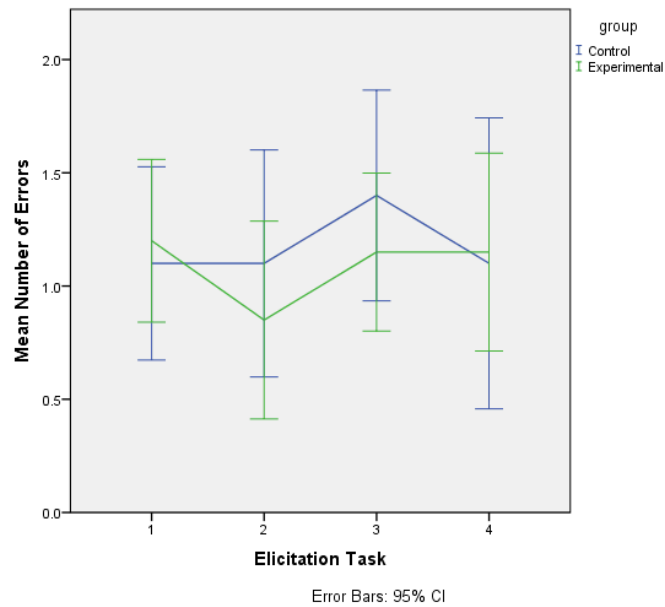| Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	1.20	0.77	0.85	0.93	1.15	0.75	1.15	0.93
Control	19	1.10	0.91	1.10	1.07	1.40	0.99	1.10	1.37

Figure 7.18: Effectiveness of WCF on Masculine/Feminine Noun Morphology Errors over Time

142

These findings are interesting because they contradict the research hypothesis. However, there is an explanation for the apparent contradiction. We will go into more detail about these particular findings in the Chapter 8 discussion, but for now, it seems clear that another stipulation is necessary if WCF is not to be rendered ineffective: *Feedback provided on one occurrence of a specific word does not lead to knowledge of other words*. That is, providing the *mf* symbol over the Spanish word "teatra" because the correct usage is "teatro" (theatre) may provide knowledge to the L2-Spanish student on how to write this particular word in the future. However, it does not provide sufficient declarative knowledge on correct gender selection for all Spanish words. The L2-Spanish student will not glean gender knowledge about the Spanish word "piscina" (pool) because he or she received feedback on the Spanish word "teatro." We see similar results with the mechanical subcategory of spelling in this study. Getting feedback on a particular word is an idiosyncratic piece of knowledge that does not apply to other words, and the *sp* symbol was ineffective in improving experimental group participants' overall spelling error rates relative to the control group participants' spelling ($p = 0.857$). This will be discussed in more detail in Chapter 8.

Subject-Verb Agreement Errors. Now for an examination of the effect of the indirect WCF symbol *s-v* on improving participants' error rates on subject-verb agreement errors. Remember, findings on this construction are important to the current study because subject-verb agreement was amenable to positive change in English in Wagner (2013), and participants demonstrate significantly better uptake behavior for this construction compared to the non-binary option constructions in the L2-English WCF

Uptake Study. It was predicted that the increased complexity of subject-verb agreement in Spanish would render the WCF symbol *s-v* ineffective for improving experimental group participants' subject-verb error reduction rates compared to the control group participants'. Findings confirm this prediction. Results for subject-verb agreement errors do not show a statistically significant difference in the reduction of subject-verb errors over the four elicitation tasks between participants from the experimental group and participants from the control group (F = 0.260, p = 0.854).

Figure 7.19: Subject-Verb Agreement Errors
Descriptive statistics for numerical changes across elicitation tasks

Once again, a quick glance at the graph demonstrates that participants behave similarly with regards to their error reduction tendencies for this grammatical category.

Both experimental and control group participants either reduce subject-verb agreement

errors marginally or actually increase their errors across the elicitation tasks. Participants

in the experimental group reduce their subject-verb agreement errors from 41 to 36

between tasks 1 and 2 (-13%); however, they increase errors from 41 to 50 between tasks

1 and 3 (+22%) and from 41 to 48 between tasks 1 and 4 (+17%). Control participants

increase their subject-verb agreement errors from 37 to 38 between tasks 1 and 2 (+2%),

from 37 to 41 between tasks 1 and 3 and between tasks 1 and 4 (+11%). Based on this

data, it appears the control group actually outperformed the experimental group in this

category.

Table 7.10 compares mean scores for total errors between experimental and

control groups across the four elicitation tasks for subject-verb agreement, and Figure

7.20 shows the graphic representation of the values in Table 7.10.

Table 7.10: Subject-Verb Agreement Errors
Descriptive statistics for mean test scores by group and elicitation task

Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	1.80	1.51	1.60	2.35	2.30	1.69	2.15	1.46
Control	19	2.10	1.68	2.10	2.31	2.25	1.65	2.30	2.36
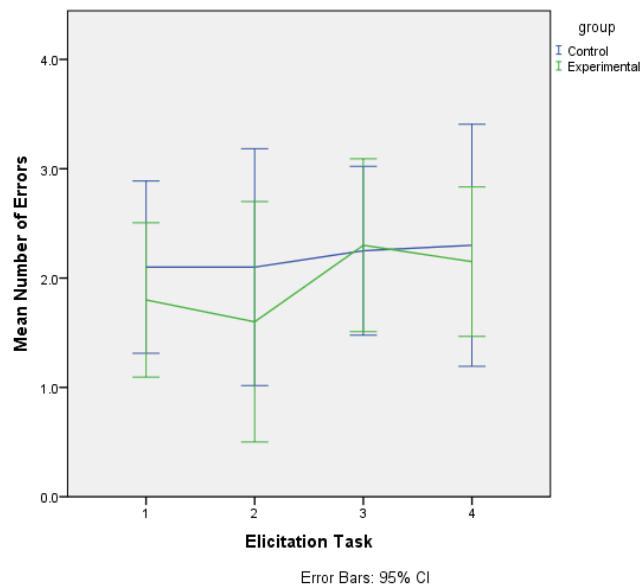
Figure 7.20: Effectiveness of WCF on Subject-Verb Agreement Errors over Time

While the indirect WCF symbol *s-v* is effective at reducing errors for L2-English students, it is ineffective at reducing errors for L2-Spanish students. This is a result of the increased complexity of subject-verb agreement in the morphologically rich language of Spanish. These findings confirm the binary vs. n-ary (n > 2) hypothesis.

Verb Tense Errors. As predicted, verb tense errors in Spanish are not amenable to positive change via the use of the indirect WCF symbol *vt*. Verb tenses in both English and Spanish are extremely complex, and providing a symbol that points out verb tense errors does not provide sufficient declarative knowledge for correct usage in new writing samples. Results for verb tense errors do not show a statistically significant difference in the reduction of verb tense errors over the four elicitation tasks between

146

participants from the experimental group and participants from the control group (F = 0.074, p = 0.974).
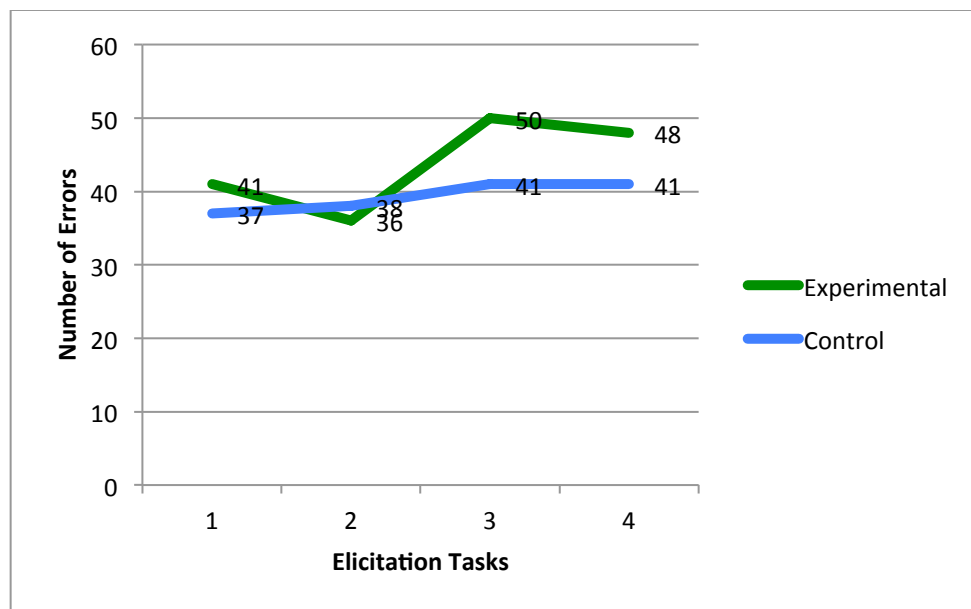

Figure 7.21: Verb Tense Agreement Errors
Descriptive statistics for numerical changes across elicitation tasks

Participants in the experimental group make only marginal reductions to their verb tense errors from 25 to 24 between tasks 1 and 2 (-13%) and from 25 to 19 between tasks 1 and 3 (-22%). Then, they increase verb tense errors from 25 to 31 between tasks 1 and 4 (+13%). Control participants follow a very similar pattern in error reductions. They decrease their verb tense errors from 23 to 20 between tasks 1 and 2 (-4%) and from 23 to 18 between tasks 1 and 3 (-24%). Then they increase verb tense errors from 23 to 26 between tasks 1 to 4 (+24%).

Table 7.11 compares mean scores for total errors between experimental and control groups across the four elicitation tasks for verb tense, and Figure 7.22 shows the graphic representation of the values in Table 7.11.

Table 7.11: Verb Tense Agreement Errors
Descriptive statistics for mean test scores by group and elicitation task

Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	1.05	0.89	0.85	1.39	0.75	0.64	1.15	0.99
Control	19	1.35	1.66	1.35	1.50	1.10	1.37	1.70	2.11
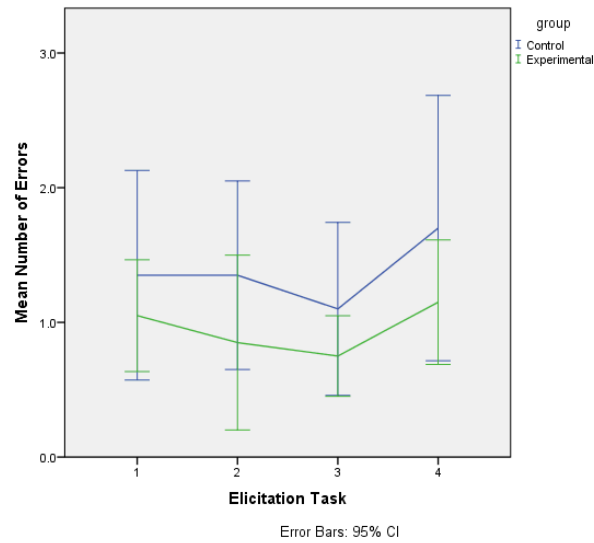
Figure 7.22: Effectiveness of WCF on Verb Tense Errors over Time

Findings on the effectiveness of the *vt* symbol help to confirm the binary vs. n-ary (n > 2) hypothesis. Correct verb tense usage in Spanish is dependent upon a large variety of options, and the *vt* symbol cannot clarify the correct usage of so many forms and form-meaning relationships.

Article Errors. That brings us to an examination of WCF effectiveness on the final grammatical category articles. Findings for article error reductions confirm predictions made at the start of the study based on the binary vs. n-ary (n > 2) hypothesis. Experimental group participants do not outperform control groups in their error reduction tendencies over the four elicitation tasks. Results for article errors do not show a statistically significant difference in the reduction of article errors over the four elicitation tasks between participants from the experimental group and participants from the control group ($F = 1.145$, $p = 0.334$).
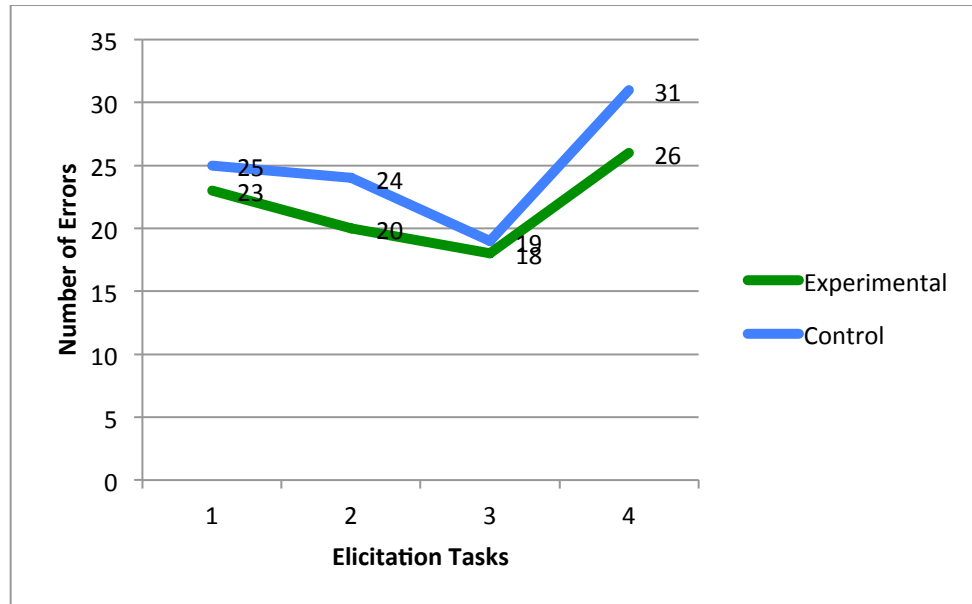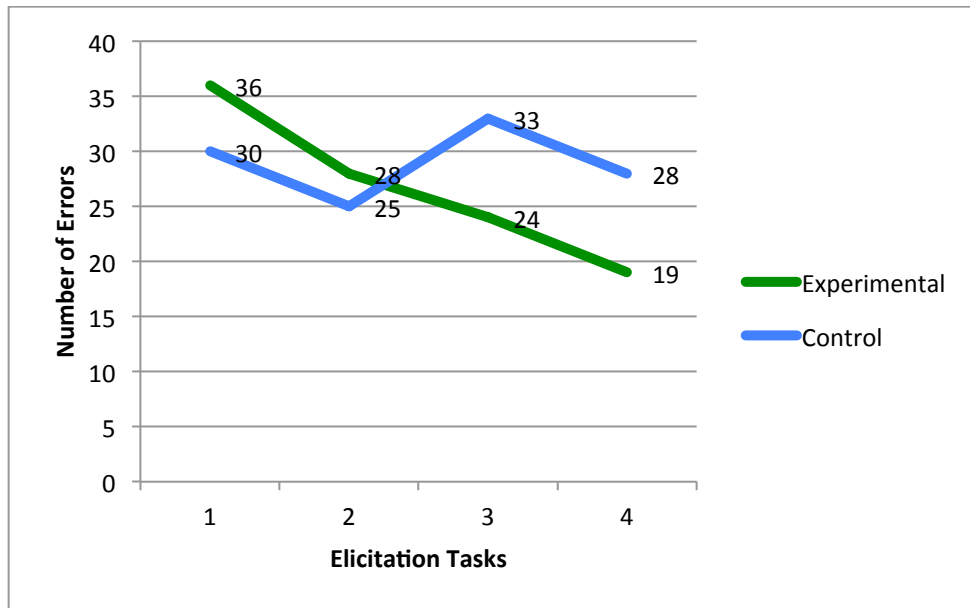
Figure 7.23: Article Errors
Descriptive statistics for numerical changes across elicitation tasks

Participants in the experimental group reduce their article errors from 36 to 28 between tasks 1 and 2 (0%), from 36 to 24 between tasks 1 and 3 (-%), and from 36 to 19 between tasks 1 and 4 (-%). Control participants decrease their article errors from 30 to 25 between tasks 1 and 2 (-17%), increase their article errors from 30 to 33 between tasks 1 and 3 (+10%), and decrease their errors from 30 to 28 between tasks 1 to 4 (-7%).

Table 7.12 compares mean scores for total errors between experimental and control groups across the four elicitation tasks for articles, and Figure 7.24 shows the graphic representation of the values in Table 7.12.

Table 7.12: Article Errors
Descriptive statistics for mean test scores by group and elicitation task

Group	N	Elicitation 1		Elicitation 2		Elicitation 3		Elicitation 4	
		M	SD	M	SD	M	SD	M	SD
Experimental	21	1.70	0.98	1.30	1.30	1.00	0.86	0.95	0.89
Control	19	1.60	1.64	1.35	1.18	1.85	1.63	1.40	1.79
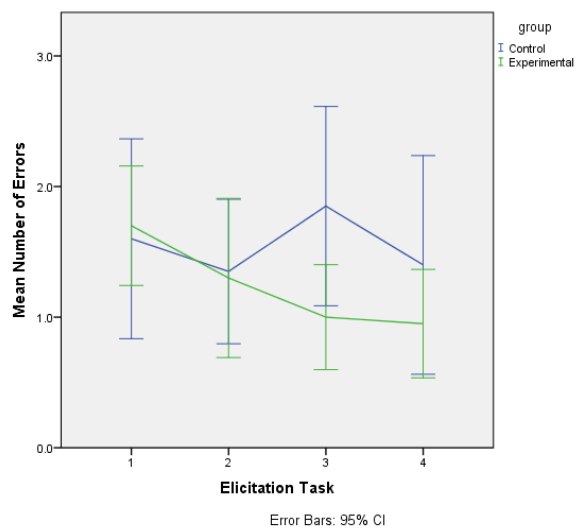
Figure 7.24: Effectiveness of WCF on Article Errors over Time

Although descriptive statistics appear to show a difference in experimental and control group participants' error reduction tendencies, this difference does not reach a level oP Valueal significance.   This is precisely what we could expect based on the research hypothesis.

151

Findings, Research Questions, and the Research Hypothesis

Overall, findings from the L2-Spanish WCF Study provide the following

answers to the studies research questions:

> **RQ 1**: Does WCF and the correction process aid in the transition from
> declarative to proceduralized knowledge of specific grammatical constructions in
> the writing modality? And if so, for which constructions is this transition
> observed?

Findings generally indicate that WCF does aid in the transition process from declarative

to proceduralized knowledge for singular/plural noun phrase agreement, overall

masculine/feminine usage, and masculine/feminine noun phrase agreement. For all the

other categories, the one-way repeated ANOVA analysis of the data demonstrates that

WCF does not aid in the transition process.

> **RQ 2**: Given the answers to RQ1, is there a grammatical complexity threshold
> beyond which WCF will no longer be operative in supplying the declarative
> knowledge necessary for L2-writing students to practice and eventually
> proceduralize?

Findings confirm a grammatical complexity threshold beyond which WCF will no

longer be operative. Experimental group participants outperform experimental groups on

all binary option constructions except overall singular/plural error reductions and binary

option rules applied idiosyncratically.

Therefore, the binary vs. n-ary (n > 2) is upheld with the caveat regarding

idiosyncratic applications of binary rules:

> **Hypothesis**. WCF is only effective in imparting declarative knowledge thereby
> increasing linguistic accuracy (proceduralization) for grammatical constructions
> that include a binary choice of options for correct usage (binary vs. n-ary (n >
> 2)).

Further discussion of these results in combination with results from the L2-English WCF

Uptake Study can be in the discussion section in Chapter 8 below.

CHAPTER EIGHT: DISCUSSION OF FINDINGS

Main Findings

Generally speaking, the binary vs. n-ary (n > 2) hypothesis is confirmed by the findings herein. The binary nature of indirect WCF (i.e., Constructions that are not marked are correct, and constructions that are marked are incorrect.) makes it operative for binary constructions (i.e., the other option is correct). Providing WCF on binary forms is equivalent to a teacher correcting a true/false test and then asking students to revise their answers. If a "T" is marked wrong, then students know with 100% certainty that "F" is the correct answer and can revise their answers accordingly. This type of certainty in WCF error correction equates to meaningful practice (i.e., strong uptake) and eventual proceduralization of rules. However, providing WCF on non-binary forms is equivalent to a teacher correcting a multiple choice test by simply marking "A" wrong and asking students to revise their answers. The students still must choose from B, C, D, or E. They cannot be sure they have chosen correctly even when they have revised their errors correctly. Findings here demonstrate that, under such circumstances, uptake suffers and that meaningful practice cannot occur to aid in the transition from declarative to proceduralized knowledge.

The most significant finding of the study is that WCF effectiveness has a complexity threshold of binary, which is an extremely limited range of effectiveness.

Results indicate that WCF is effective in helping L2 students acquire grammatical systems at or below this threshold. However, WCF is not effective in helping L2 students acquire grammatical structures that are inherently more complex than this threshold (i.e. include more than a binary option for correct usage). Indeed, if a construction has a binary rule to apply in a specific syntactic location (i.e., the end of a noun for singular/plural and at the end of a verb for third-person singular subject-verb agreement) and the rule can apply generally to all cases in that location, WCF is can impart sufficient declarative knowledge for acquisition of the whole grammatical paradigm to begin. Binary option structures are corrected a majority of the time in the L2-English study, and binary option constructions demonstrate evidence of proceduralization in the L2-Spanish study.

Findings on Spanish masculine/feminine grammar provide an interesting proviso regarding the use of indirect WCF on binary forms. While participants in the L2-Spanish study demonstrate proceduralization of overall masculine/feminine knowledge as well as masculine/feminine noun phrase agreement knowledge, they do not demonstrate proceduralization for the gender selection of individual words. Feedback on one word increases a student's declarative knowledge of that particular word's gender assignment, but it does not increase their declarative knowledge for the gender selection of all Spanish words. In this case, the location for morphological selection is always at the end of a noun (like singular/plural), but the correction does not apply generally to all instances of all nouns (i.e., some are masculine and some are feminine). This could be the reason for why L2-Spanishs participants from the experimental group did not significantly

155

outperform participants from the control group in their reduction of masculine/feminine errors (i.e., gender selection) on individual words.

However, this does not mean that the WCF *mf* symbol is ineffective at providing declarative knowledge for gender selection of specific words. Progress on these idiosyncratic items will be item-by-item. Indeed, no matter what correction methodology is used, progress on idiosyncratic items will, by the very nature of idiosyncrasy, be item-by-item progress. Indirect WCF is no worse than any other instructional strategy and can be used to deal with idiosyncratic, binary constructions one at a time (e.g., gender selection for one Spanish noun at a time). Consequently, indirect WCF is effective on binary constructions and provides limited guidance only for binary, idiosyncratic paradigms.

The *mf* symbol is effective in drawing students' attention to the noun phrase and its boundaries, and the repetition of the *mf* feedback symbol across noun phrases imparts declarative knowledge of agreement for that syntactic location. Declarative knowledge of the binary nature of gender assignment in Spanish that makes all noun phrase materials agree as either masculine or feminine (even where students select the wrong gender) can be imparted, practiced, and proceduralized.

The L2-Spanish study demonstrates a binary ceiling on WCF effectiveness. Whereas participants demonstrate significant gains in singular/plural usage and masculine/feminine agreement across noun phrases, they show no significant accuracy increases and no proceduralization of non-binary constructions. Therefore, the predictive utility of the binary vs. n-ary (n > 2) hypothesis has been upheld with the

156

stipulation that even binary construction rules cannot be idiosyncratic for WCF to be effective on an entire grammatical category (e.g., improvement to correct gender selection for all Spanish nouns). However, indirect WCF can provide some limited guidance on binary, idiosyncratic usage (i.e., one item at a time). More important, the binary vs. n-ary (n > 2) hypothesis and SAT analyses can continue to act as a foundation for further study of not only WCF and varied types of WCF, but also other L2-teaching strategies that attempt to impart knowledge and increase L2 students' grammatical competence.

In addition to the main finding of a complexity threshold, data and statistical analyses augment the strong-interface position of SLA and data and statistical analyses demonstrate that SAT (a theory grounded in the strong-interface position) has a great deal of explanatory and predictive power when it comes to analyzing and assessing L2-teaching strategies like WCF.   When sufficient declarative knowledge of L2 grammatical constructions is imparted from L2-teacher to L2-student and when that knowledge can be practiced sufficiently, correctly, and meaningfully; that knowledge can transition from declarative to proceduralized knowledge.

The L2-English study confirms that L2-English learners understand English subject-verb agreement and singular/plural WCF symbols and the underlying rules they represent. This understanding is demonstrated by participants' overwhelming attempts to fix these two error types and their abilities to revise these errors correctly in final drafts a vast majority of the time (79%). The exact opposite is true of participant behavior regarding correction attempts and correct revisions of the non-binary option constructions

(49%) in the L2-English study. They do not gain sufficient declarative knowledge from the WCF symbols for verb tense, article usage, or preposition usage. Practice is therefore inconsequential, and proceduralization will not be attained for these constructions (as further confirmed by the L2-Spanish WCF Study). Where WCF fails to provide sufficient declarative knowledge of a construction, it sabotages the acquisition process from the start.

Because participants revised some constructions significantly better than others in the L2-English study and proceduralized some constructions significantly better than others in the L2-Spanish study, the explanatory power of the non-interface position is weakened. If no connection exists between conscious knowledge of forms and an eventual proceduralization and automatization of said forms, participants should have failed to significantly revise *all error types* in the L2-English study. They should also have failed to demonstrate proceduralization of *all error types* in the L2-Spanish study. This was not the case.

Of course, L2 learners may be able to figure out declarative knowledge eventually through various means in order to move forward with acquisition, even if this knowledge is beyond the threshold of WCF effectiveness. However, findings here demonstrate that WCF was of no benefit beyond its effective range. In cases of constructions beyond its threshold of effectiveness, WCF is not accomplishing anything.

Findings and Grammatical Complexity/Difficulty

A linguistic approach to the study of WCF leads almost immediately to a linguistic analysis of grammatical complexity. Findings support the binary vs. n-ary (n > 2) hypothesis and provide a strong and accurate approach for measuring grammatical complexity. This approach to defining grammatical complexity combines two separate approaches (i.e., linguistic definitions and pedagogical definitions) and provides us with the opportunity to hypothesize and falsify claims of simplicity and complexity with empirical evidence.

This study defines grammatical complexity very narrowly and according to a strict and easily falsifiable line of demarcation separating simple from complex. This approach to defining grammatical complexity is grounded in one of the linguistic definitions of complexity, namely the inherent complexity of structures (Hulstijn & de Graaff, 1994; Givon, 1991; 1995; Housen, Pierrard, & Van Daele, 2005; Robinson, 1996; DeKeyser, 1998; Doughty & Williams, 1998; Spada & Tomita, 2010). While the researchers listed above agree that structures are either inherently simple or complex, "inherent complexity" is a broad term that can be interpreted in a variety of ways. Also, this inherent complexity can be measured in a number of ways according ot a large variety of criteria.

As mentioned in the literature review in Chapter 2, objectively assessing grammatical complexity can be extremely difficult. As researchers apply different criteria of complexity to varied constructions, inevitably miscalculations are made regarding constructions' levels of complexity. Hulstijn and de Graff (1994) argue that "The degree of complexity is contingent not so much on the number of forms in a

paradigm, but rather, on the number (and/or the type) of criteria to be applied in order to arrive at the correct form" (p. 103). Their conclusion is that complexity is dependent upon the number of transformations a particular construction undergoes. While this is an excellent hypothesis that can be easily tested, the criteria applies only to constructions that have transformations and discounts all other reasons for why constructions that do not have transformations might be complex. Based on their complexity criteria, Spada and Tomita (2010) provide following predictions for what should be simple and complex in English:

Table 8.1: Hulstijn and de Graaf's Complexity Predictions

Simple	Complex
Tense Articles Plurals Prepositions Subject-verb Inversion Possessive Determiners Participial Adjectives	Dative Alternation Question Formation Relativization Passives Pseudo-cleft Sentences

Of course, we know that correct tense, article usage, and preposition usage come late in L2-English acquisition and often not at all. Findings from the current study confirm that these are complex constructions, especially relative to English subject-verb agreement and singular/plural constructions. Clearly, applying one criterion across all grammatical

160

constructions cannot possibly distinguish which structures are complex and which are not.

Also remember that applying different criteria to the same structure can lead to disagreements over a construction's complexity. Krashen (1982) claims English subject-verb agreement is simple based on the number of alternative forms (add –s or do not add –s to the base form of a verb), Ellis R. (1990) claims it is complex based on the construction's processing demands (the long distance between the subject and the end of the verb, and DeKeyser (1998) claims that the construction is complex because of its highly syncretic nature (the combination of three abstract concepts— present tense, third person, singular). Findings here demonstrate that subject-verb agreement is a simple construction in English, and Krashen's approach to defining complexity appears to be most accurate.

Based on this studies confirmation of the binary vs. n-ary (n > 2) hypothesis, an approach to defining complexity based on the number of alternative forms seems to provide an accurate account of what should be defined as simple and what should be defined as complex. DeKeyser (2005) defines complexity by counting the number of forms (i.e., linguistic substance), the number of meanings, and the number of form-meaning relationships inherent to given constructions. This makes sense intuitively. Adding or not adding "s" to a noun to make it plural or singular is simple. Generally speaking, adding "s" means more than one and not adding "s" means one. This one-to-one relationship is easy to learn and apply without making too many mistakes. Knowing how to correctly use the English verb tense system, which has twelve tenses, each of

which has a lot of linguistic materials (e.g., will not have been going) and conveys subtle distinctions of time and aspect is far more complex. Moreover, some tenses can be used to convey more than one meaning. For example, the simple present can be used to convey actions done repeatedly or actions in the future:

1. I go to school every day.

2. I go to school next year.

This represents a form-meaning relationship that is not a one-to-one relationship, and this, according to DeKeyser, creates opacity and increases complexity. That is, where different forms stand for the same meaning and the same form stands for different meanings, opacity increases, and this increases complexity.

It is extremely important to keep opacity in mind when evaluating a constructions' level of complexity. Take English articles for example. At first glance, one might predict that English article usage would be the ideal construction to test the binary vs. n-ary (n > 2) hypothesis to see if the boundary could be moved to binary vs. n-ary (n > 3). Correct article usage in English requires the correct use three options: the definite, the indefinite, and the null article. However, opacity makes articles far more complex. They do not have one-to-one form-meaning relationships.

Regardless, findings from the current study confirm a binary ceiling for what is simple, and this ceiling simply divides constructions into two groups, simple and complex. Of course, we know that there is probably a scale of complexity, and findings here provide a foundation for establishing such a scale. The best way to determine if there is a higher threshold of treatability beyond binary option constructions is to use the

same intervention, indirect WCF, to examine a construction in a given language with three options for correct usage. Crucially, this construction's three options must have one-to-one meaning relationships. Based on findings here and the nature of indirect WCF, it can be predicted that such a construction would not be amenable to positive change. L2 students encountering the symbol for such an error would have to guess between the other two options. After making the guess, they would not know if they guessed correctly. The fact that they have to guess demonstrates that the feedback has not provided sufficient declarative knowledge and guessing would not constitute meaningful practice.

This demonstrates a pedagogical approach to defining grammatical difficulty. After hypothesizing the level of complexity associated with a given construction, simple or complex, the same pedagogical instrument can be used to test the complexity of the construction. Remember that pedagogical approaches to defining complexity are explained in terms of students' abilities to understand or learn different grammatical constructions. This also makes intuitive sense. After all, what are we saying when we say a construction is simple or complex? We are saying that it is either easy or difficult for L2 learners to use correctly and/or acquire accurately.

Usually this approach is characterized by teachers identifying difficult constructions based on their observations of L2 students' performance or errors that L2 students make regularly (Robinson, 1996) or based on students' assessments of complexity compared to their performance (Scheffler, 2009). The approach used in this study is more objective. The binary ceiling sets the hypothesized ceiling. Then, we

163

determine if that ceiling is an accurate gauge of complexity by pointing an error and seeing if L2-students can repair the error on their own (i.e., successful uptake) and if the L2-students can demonstrate increased accuracy for the construction over time (i.e., proceduralization). If pointing to an error is sufficient for making an L2 student understand the underlying rules of a construction and if the practice of correction of these errors leads to greater accuracy, we can label said construction as simple.

This combined linguistic/pedagogical approach to defining grammatical complexity may not be perfect. Undoubtedly, some will argue that two dependent variables are being examined, the effectiveness of indirect WCF and grammatical complexity. However, applying the same intervention to all constructions can help us to delineate simple from complex. Also, setting a binary ceiling based on the number of forms and form-meaning relationships with in a construction provides a falsifiable claim that others can test.

Most importantly, this approach does not suffer from the problems associated with applying one subjective criterion of complexity to all constructions (e.g., Hulstijn and de Graaf) or many different subjective criteria to the same construction (e.g., Krashen, Ellis R., and DeKeyser). It also is not dependent on circular logic of developmental readiness definitions of complexity: The construction is difficult because it is learned later, and it is learned later because it is difficult.

Furthermore, this approach to defining grammatical complexity could be tested without the use indirect WCF. L2 teachers could sit with L2 students and point out errors in their writing to see if the students could correct the error correctly. If the student

successfully corrects the error without any other guidance a majority of the time (like we saw in the L2-English Study), we could assume the construction is relatively simple. These constructions could be labeled as "simple to teach/learn/acquire" and simple interventions like indirect WCF could be administered to aid in students' acquisition of these constructions.

Findings on grammatical complexity herein are also relevant to Ferris' treatable/non-treatable distinction. Ferris claims that this distinction should be based on whether or not grammatical constructions are rule based or idiosyncratic. This is at least partially confirmed by findings that demonstrate a lack of preposition corrections in the L2-English study. Preposition usage is extremely idiosyncratic, and the WCF *pr* symbol could not elucidate rules for their idiosyncratic usage. Of course, *pr* also could not elucidate preposition rules because there are so many prepositions to choose from when making a correction. However, findings here do not support the claim that all rule-based systems should be treatable with the use of WCF. WCF effectiveness is not determined by whether a system has rules or not. It is determined on the level of complexity and the number of rules within the systems, something Ferris does not mention in her discussions of treatability. The following table shows the differences between Ferris's treatability distinctions for English and the treatability findings of the L2 English WCF Uptake Study:

Table 8.2: A Comparison of Ferris's Treatability Distinctions for English versus The Current Study's Findings

Ferris		L2-English WCF Uptake Study	
Treatable	**Non-treatable**	**Treatable**	**Non-treatable**
Singular/Plural	Prepositions	Singular/Plural	Verb Tense
Subject-Verb Agreement		Subject-Verb Agreement	Articles
Verb Tense			Prepositions
Articles			

WCF is effective at explicating simple systems but not complex ones. Rules that are explained sufficiently become declarative knowledge that can be practiced and proceduralized. L2-English study participants demonstrated declarative knowledge of simple, binary rules in their uptake results. Rules that are not explained sufficiently are not practiced or are practiced incorrectly and are not proceduralized over time. Participants were unable to demonstrate uptake of the non-binary constructions.

Ferris's treatability distinctions based on her "rule-based" analysis does not appear to be accurate based on the L2-Spanish WCF Study either:

Table 8.3: A Comparison of Ferris's Treatability Distinctions for Spanish versus The Current Study's Findings

Ferris		L2-Spanish WCF Study	
Treatable	**Non-treatable**	**Treatable**	**Non-treatable**
Singular/Plural Masculine/Feminine Agreement Subject-Verb Agreement Verb Tense Articles		Singular/Plural Masculine/Feminine Agreement across the NP Masculine/Feminine Selection for Individual Words (One-by-One)	Subject-Verb Agreement Verb Tense Articles

Clearly, judging treatability, and by extension WCF effectiveness, exclusively on whether or not a construction is rule based or not will not provide the same predictive utility as judging treatability on the complexity of these rules.

Findings and Pedagogical Implications

Findings from this study provide detailed explanations for the apparent contradictory theoretical interpretations of WCF usefulness and empirical findings of WCF effectiveness in the literature. This study demonstrates that an all-or-nothing approach to examining WCF effectiveness cannot provide a clear understanding of how and why WCF actually works. An SAT analysis of WCF helps to clarify many of the

apparent contradictions in the literature between researchers on both sides of the WCF debate.

Truscott's (1996) opposition to WCF is partly justified by my findings, since WCF is ineffective for some constructions. However, this opposition is too wide ranging, since for many constructions WCF provides sufficient declarative knowledge to allow for acquisition to occur. Certainly, Truscott is correct that L2 acquisition of many grammatical constructions and linguistic features is the result of L2 input and unconscious or implicit learning. However, results from this study indicate that instructed learning, like the type that occurs with the correct use of WCF, can aid in long-term acquisition of grammatical knowledge and the eventual automatized use of such knowledge.

SAT analyses of WCF not only provide a way to quantify effectiveness in WCF research but also provide a paradigm for L2-writing instructors to assess WCF teaching strategies as well as other L2-teaching strategies that attempt to improve L2 students' accuracy. SAT is suitable for WCF analysis for two reasons: (1) the current study confirms a connection between declarative and proceduralized knowledge can exist with a few stipulations regarding complexity and (2) the tenets of SAT mirror how grammar instruction is often expected to work by L2 instructors.

Based on this study's findings, teachers should use the following diagnostic to determine if and when WCF should be used to improve L2 students' grammatical accuracy:

1. Is the construction a binary choice? If no, stop. If yes, continue.

2. Is the construction stable or idiosyncratic? If idiosyncratic, continue, but anticipate only limited, item-by-item improvement.

This diagnostic correctly predicts which grammatical constructions should be amenable to change and which should not. Also, L2-writing teachers might decide that using indirect WCF on binary, idiosyncratic forms like masculine/feminine gender selection for individual nouns in Spanish is worthwhile. As mentioned above, they could affect positive change for the correct gender selection of a great number of individual nouns over the course of a semester-long writing class, even if they cannot hope to improve accuracy to the category overall.

Finally, once it has been determined that a given type of WCF can impart the necessary declarative knowledge for meaningful practice to begin (i.e., binary and not idiosyncratic), a failure of WCF to significantly improve accuracy rates should be a result of failures in the practice procedure, and this can be examined separately in future studies of WCF.

Suggestions for Future Research

SAT analyses of WCF across studies can quantify effectiveness and make results more comparable across studies. WCF success can be gauged by its ability to (1) provide sufficient declarative knowledge and (2) provide sufficient practice procedures once it has been confirmed that this declarative knowledge has been imparted. WCF failure points to a lack in one of these two stipulations. Consequently, WCF that is ineffective can be analyzed to determine if it can be modified to impart sufficient declarative knowledge (e.g., Feedback is not working as expected.), and/or the practice procedure

can be modified/improved to provide sufficient practice (e.g., More iterations of practice are required.).

Future studies could further examine WCF's success on binary forms. It might be possible to break some complex constructions down to binary options for future examinations of WCF effectiveness and to determine objectively assess grammatical complexity. Such examinations could focus on choices between two correct usage rules instead of all the rules that underlie a construction. For example, in the current study, the symbol *vt* was expected to impart a large amount of declarative knowledge (i.e. syntactic and sematic rules of 12 verb tenses). However, if this *vt* symbol was used to distinguish between only the simple past and the simple present, for example, for writing samples based on writing prompts that require students to use either one or the other tense, we might see significant gains in their declarative knowledge and proceduralization of rules for these two tenses. If experimental groups demonstrate increased accuracy and proceduralization of the two tenses over time relative to control groups, the use of WCF for addressing binary grammatical choices would be strengthened. WCF might also be examined to determine its effectiveness for increasing accuracy on the usage of two prepositions (e.g., in or on).

At first glance, this may appear to be an endorsement of the SLA "article studies" that examined WCF effectiveness for increasing L2-English students' procedural knowledge of only two functional uses of the article system. This is not the case. Using WCF to impart declarative knowledge of sub rules of the article system is not only impractical, but it also imparts rules that run counter to actual English article usage. Here

the feedback is imparting a piece of knowledge that is binary (i.e., choose the definite article here and the indefinite here), but both the definite and indefinite article also correspond to other semantic and syntactic uses and meanings. This lack of a one-to-one correspondence is precisely what makes acquisition of English articles so difficult. Instead of examining WCF and its effects on sub rules of the article system, SLA researchers should be applying theories of acquisition that provide explanations for how and why acquisition of some structures can be positively influenced in instructed environments and others cannot.

The successful SAT analysis utilized in the current study could also be applied to comparative studies of different types of direct and indirect WCF. Instead of simply claiming one works better than another, the different types of feedback could be assessed to determine which one provides more declarative knowledge (i.e., examine uptake) and which leads to stronger increases in proceduralization of this knowledge (i.e., examine accuracy increases between experimental and control groups). The findings herein demonstrate that indirect WCF correction symbols are effective on binary forms. As mentioned previously, providing these symbols for binary forms is tantamount to providing the correct revision (i.e., This one is wrong; the correction is the only other option.). While some may argue that direct feedback does not promote problem solving and engagement in the practice procedure, the two types of feedback appear to provide declarative knowledge in very similar ways.

Based on the current studies' findings the following research hypothesis could be examined:

> Direct feedback imparts more declarative knowledge than indirect feedback, and this increased conveyance of declarative knowledge makes the practice procedure more beneficial (as demonstrated by increases in uptake), and this leads to increases in evidence of proceduralization (as demonstrated by accuracy increases for those who receive direct WCF relative to those who receive indirect WCF).

Bitchener and Knock (2010) claim that practitioners of direct WCF believe it "provides more information on complex errors" (p. 210). Based on the findings from the current study, it must provide more information than a symbol, circle, or number at the end of a line of text (i.e., indirect WCF types). They also claim that practitioners believe direct feedback "provides more explicit feedback on hypotheses that may have been made" (p. 210). Testing the hypothesis above could help to confirm or deny these claims. Also, if an "uptake" study was conducted and demonstrated similar uptake tendencies on binary constructions for both indirect and direct WCF, the practice procedure could also be examined to determine if engagement and problem solving play an important role in the correction process.

Conclusion

WCF effectiveness can be better understood under SAT analyses of how the intervention actually works. This dissertation demonstrates that it is the level of grammatical complexity that determines whether or not WCF will be useful in helping L2 students acquire grammatical constructions. Specifically, the WCF interventions examined in this study should only be expected to help in the acquisition of whole grammatical paradigms if said structures include a binary rule that is not idiosyncratically applied. If the binary construction rules apply idiosyncratically, WCF can still be

172

effective but only one item at a time. If the construction has more than a binary option

for correct usage, WCF will not help in the acquisition process. This is a result of WCF

not being able to impart more than binary declarative knowledge.

The L2-English Study demonstrates that WCF imparts declarative knowledge of

simple, binary structures but not non-binary structures. This study examined WCF

effectiveness at the beginning of the process of acquisition outlined by SAT. That is, it

focused explicitly on WCF symbols' capacities to impart clear and useful declarative

knowledge that could be practiced meaningfully. By examining and highlighting

participants different correction/practice behaviors when confronted with WCF for both

binary and non-binary constructions, we were able to demonstrate that WCF symbols

effectively impart knowledge of binary forms and that WCF symbols do not effectively

impart knowledge of non-binary forms. Indeed, when sufficient knowledge is not

imparted, the acquisition process breaks down from the beginning. Meaningful practice

of knowledge cannot occur. Findings from the L2-English WCF Uptake Study

demonstrate that when the SAT prerequisite of "imparting adequate and clear declarative

knowledge" is not met, the SAT process of acquisition cannot even begin.

The L2-Spanish Study examines WCF effectiveness across the whole process of

acquisition outlined by SAT (i.e., impart declarative knowledge, provide sufficient

practice, and declarative knowledge transitions to proceduralized knowledge). Generally

speaking, it demonstrates that constructions designated as complex by the binary vs. n-

ary (n > 2) research hypothesis are not proceduralized via WCF and correction practice,

while constructions designated as simple show evidence of proceduralization via

accuracy increases over the elicitation period. This study demonstrates that WCF imparts knowledge of binary structures but not non-binary structures. It also demonstrates that the practice procedure of correcting errors over the elicitation period is sufficient for a transition of knowledge for whole grammatical paradigms that include binary, non-idiosyncratic options for correct usage.

Findings from the two studies herein are complimentary. The L2-Engish WCF Uptake Study demonstrates strong uptake for binary option constructions and weak uptake for non-binary constructions. The L2-Spanish WCF Study generally demonstrates proceduralization for binary option constructions and a lack of proceduralization for non-binary constructions. These findings, taken together, demonstrate the WCF should be used more economically in L2 classrooms to help aid the acquisition process for constructions for which it can actually impart sufficient declarative knowledge.

APPENDIX

A. **Recruitment Script L2-English WCF "Uptake Study**

Mr. Wagner is doing a study for his doctorate. He is studying to see how students learn to write better in their second language. You will complete all the assignments as regular classwork, but if you choose to be in the study, your writing data will be included in Mr. Wagner's study. Your name will not be included. If you choose not to participate, your writing data will not be included in Mr. Wagner's study. You do not have to participate in this study if you do not want to participate. You can just tell me, "I do not want to be in the study." If you decide to participate now, but later decide that you do not want to be in the study, you can tell me, "I do not want to be in the study anymore." This will not affect your grade in Mr. Wagner's class. You can change your mind at any time. You will not be penalized for not participating or for ending your participation.

If you are 18 or older and want to have your data used in Mr. Wagner's study, you will sign the form entitled "INFORMED CONSENT FORM (FOR 18+ Students)."

If you are under 18 and you want to participate in Mr. Wagner's study, you will sign the form entitled "INFORMED ASSENT FORM (FOR STUDENTS)". Your parents will also have to agree to let you participate. If they agree to let you participate, they will sign the form entitled "INFORMED CONSENT FORM (FOR PARENTS)." Both of these forms must be returned by Friday of this week. If you do not return the forms by then, I will assume that you are not taking part in Mr. Wagner's study.

Mr. Wagner will not know if you are participating or not until after the semester ends and grades have been turned in.

B. **Recruitment Script L2-Spanish WCF Study**

I am doing a study for my doctorate. I am studying to see how students learn to write better in their second language. You will complete all the assignments as regular classwork, but if you choose to be in the study, your writing data will be included in my study. Your name will not be included. If you choose not to participate, your writing data will not be included in my study. You do not have to participate in this study if you do not want to participate. You can just tell me, "I do not want to be in the study." If you decide to participate now, but later decide that you do not want to be in the study, you can tell me, "I do not want to be in the study anymore." This will not affect your grade in your Spanish class. You can change your mind at any time. You will not be penalized for not participating or for ending your participation.

If you are 18 or older and want to have your data used in my study, you will sign the form entitled "INFORMED CONSENT FORM (FOR 18+ Students)."

If you are under 18 and you want to participate in Mr. Wagner's study, you will sign the form entitled "INFORMED ASSENT FORM (FOR STUDENTS)". Your parents will also have to agree to let you participate. If they agree to let you participate, they will sign the form entitled "INFORMED CONSENT FORM (FOR PARENTS)." Both of these forms must be returned by Friday of this week. If you do not return the forms by then, I will assume that you are not taking part in my study.

C. **Informed Assent/Consent Forms for L2-English WCF "Uptake" Study**

1. **INFORMED ASSENT FORM (FOR STUDENTS)**

"**Understanding the Variable Effectiveness of Written Corrective Feedback in the**

Acquisition of Grammatical Forms: A Linguistic Perspective"

RESEARCH PROCEDURES
Mr. Wagner wants to do a study to see if the way he teaches writing helps students to become better writers. If you agree to participate, you will write 3 essays. You must write these 3 essays for class even if you do not participate in the study. If you say, "**Yes**," your essays will be used in Mr. Wagner's study. If you say, "**No**," your essays will <u>not</u> be used for Mr. Wagner's study.
RISKS
There are no foreseeable risks for participating in this research.
BENEFITS:
There will be no direct benefits to participants.
Mr. Wagner's hopes his research will be used to make other ELL writing classes better so that other ELL students in other ELL classes can learn to be better writers faster.
CONFIDENTIALITY
Your name will not be used in this study. The data in this study will be confidential. Names and other identifiers will be blacked out on writing samples: (1) your name will be blacked out on all collected data; (2) a code will be placed on all collected data; (3) through the use of an identification key, the researcher will be able to link your data to your identity; and (4) only the researcher will have access to the identification key. Where names are written on writing samples, they will later be blacked out and will only be identified by participants' ID numbers.
PARTICIPATION
You do not have to participate in this study if you do not want to participate. You can just tell Ms. Godsey, "I do not want to be in the study." If you decide to participate now, but later decide that you do not want to be in the study, you can tell Ms. Godsey, "I do not want to be in the study anymore." This will not affect your grade in Mr. Wagner's class. You can change your mind at any time. You will not be penalized for not participating or for ending your participation.
CONTACT
This research is being conducted by Doug Wulf and Jason Wagner at George Mason University. You can call Doug at (703) 993-2771 and Jason at (703) 674-7194 for questions or to report a research-related problem. You may contact the George Mason University Office of Research Integrity and Assurance at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research.

This research has been reviewed according to George Mason University procedures governing your participation in this research.
* Translators will be available to students when this form is discussed in class.
CONSENT
I have read this form and agree to participate in this study.

_____   _____   _____

Name (Print)   Name (Sign)   Date
Version date: January 7, 2015

2. INFORMED CONSENT FORM (FOR 18+ Students)

"**Understanding the Variable Effectiveness of Written Corrective Feedback in the**

Acquisition of Grammatical Forms: A Linguistic Perspective"

RESEARCH PROCEDURES
Mr. Wagner wants to do a study to see if the way he teaches writing helps students to become better writers. If you decide to participate, you will write 3 essays. Your must write these 3 essays for class even if you do not participate in the study. If you say, "**Yes**," your essays will be used in Mr. Wagner's study. If you say, "**No**," your essays will <u>not</u> be used for Mr. Wagner's study.
RISKS
There are no foreseeable risks for participating in this research.
BENEFITS
There will be no direct benefits to participants.
Mr. Wagner's hopes his research will be used to make other ELL writing classes better so that other ELL students in other ELL classes can learn to be better writers faster.
CONFIDENTIALITY
Your name will not be used in this study. The data in this study will be confidential. Names and other identifiers will be blacked out on writing samples: (1) your name will be blacked out on all collected data; (2) a code will be placed on all collected data; (3) through the use of an identification key, the researcher will be able to link your data to your identity; and (4) only the researcher will have access to the identification key. Where names are written on writing samples, they will later be blacked out and will only be identified by participants' ID numbers.
PARTICIPATION
You do not have to participate in this study. Participation is completely voluntary. If you sign and return this form to school, and you also agree to participate; your essays will be used in the study. If you do <u>not</u> sign and return this form, your essays will <u>not</u> be used in this study. You can contact Ms. Shawn Godsey at sgodseyon@olatheschools.org at any time to end your participation in the study. There is no penalty for not participating or for ending participation.
CONTACT
This research is being conducted by Dr. Doug Wulf and Jason Wagner at George Mason University. You can call Doug at (703) 993-2771 and Jason at (703) 674-7194 for questions or to report a research-related problem. You may contact the George Mason University Office of Research Integrity and Assurance at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research.

This research has been reviewed according to George Mason University procedures governing your participation in this research.

CONSENT
I have read this form and agree to participate in this study.

_____   _____
Student's Name (Print)   Date

Version date: January 7, 2015

3. INFORMED CONSENT FORM (FOR PARENTS)

"**Understanding the Variable Effectiveness of Written Corrective Feedback in the**

Acquisition of Grammatical Forms: A Linguistic Perspective"

RESEARCH PROCEDURES
Mr. Wagner wants to do a study to see if the way he teaches writing helps students to
become better writers. If you let your child participate, your child will write 3 essays.
Your child must write these 3 essays for class even if he or she does not participate in the
study. If you say, "**Yes**," your child's essays will be used in Mr. Wagner's study. If you
say, "**No**," your child's essays will <u>not</u> be used for Mr. Wagner's study.
RISKS
There are no foreseeable risks for participating in this research.
BENEFITS
There will be no direct benefits to participants.
Mr. Wagner's hopes his research will be used to make other ELL writing classes better so
that other ELL students in other ELL classes can learn to be better writers faster.
CONFIDENTIALITY
Your child's name will not be used in this study. The data in this study will be
confidential. Names and other identifiers will be blacked out on writing samples: (1)
your child's name will be blacked out on all collected data; (2) a code will be placed on
all collected data; (3) through the use of an identification key, the researcher will be able
to link your child's data to his or her identity; and (4) only the researcher will have access
to the identification key. Where names are written on writing samples, they will later be
blacked out and will only be identified by participants' ID numbers.
PARTICIPATION
Your child does not have to participate in this study. Participation is completely
voluntary. If you sign and return this form to school and your child also agrees to
participate; your child's essays will be used in the study. If you do <u>not</u> sign and return
this form, his or her essays will <u>not</u> be used in this study. You can contact Ms. Shawn
Godsey at sgodseyon@olatheschools.org at any time to end your child's participation in
the study. There is no penalty for not participating or for ending participation.
CONTACT
This research is being conducted by Doug Wulf and Jason Wagner at George Mason
University. You can call Doug at (703) 993-2771 and Jason at (703) 674-7194 for
questions or to report a research-related problem. You may contact the George Mason
University Office of Research Integrity and Assurance at 703-993-4121 if you have
questions or comments regarding your child's rights as a participant in the research.

This research has been reviewed according to George Mason University procedures
governing your child's participation in this research.
* Interpretation Service: Contact Jason Wagner at jwagneron@olatheschools.org to set
up a time with a translator to discuss this form in your native language. (Por favor

contactar Jason Wagner a su jwagneron@olatheschools.org para consultar una sita con un traductor para discutir esta forma en su lengua.)

CONSENT

I have read this form and agree to participate in this study.

_____
Student's Name (Print)

_____   _____   _____
Parent's Name (Print)   Parent's Name (Sign)   Date

_____   _____   _____
Parent's Name (Print)   Parent's Name (Sign)   Date

Version date: January 7, 2015

D. **L2-Spanish Informed Assent/Consent Forms for L2-Spanish WCF Study**

1. INFORMED ASSENT FORM (FOR STUDENTS)

"Understanding the Variable Effectiveness of Written Corrective Feedback in the

Acquisition of Grammatical Forms: A Linguistic Perspective"

RESEARCH PROCEDURES
Mr. Wagner wants to do a study to see if the way he teaches writing helps students to become better writers. If you agree to participate, you will write 4 paragraphs. You must write these 4 paragraphs for class even if you do not participate in the study. If you say, "**Yes**," your paragraphs will be used in Mr. Wagner's study. If you say, "**No**," your paragraphs will <u>not</u> be used for Mr. Wagner's study.
RISKS
There are no foreseeable risks for participating in this research.
BENEFITS:
There will be no direct benefits to participants.
Mr. Wagner's hopes his research will be used to make other SLL writing classes better so that other SLL students in other SLL classes can learn to be better writers faster.
CONFIDENTIALITY
Your name will not be used in this study. The data in this study will be confidential. Names and other identifiers will be blacked out on writing samples: (1) your name will be blacked out on all the collected data; (2) a code will be placed on the collected data; (3) through the use of an identification key, the researcher will be able to link your writing samples to your identity; and (4) only the researcher will have access to the identification key. Where names are written on writing samples, they will later be blacked out and will only be identified by participants' ID numbers.
PARTICIPATION
You do not have to participate in this study if you do not want to participate. You can just tell Mr. Wagner, "I do not want to be in the study." If you decide to participate now, but later decide that you do not want to be in the study, you can tell Mr. Wagner, "I do not want to be in the study anymore." This will not affect your grade in Mr. Wagner's class. You can change your mind at any time. You will not be penalized for not participating or for ending your participation.
CONTACT
This research is being conducted by Doug Wulf and Jason Wagner at George Mason University. You can call Doug at (703) 993-2771 and Jason at (703) 674-7194 for questions or to report a research-related problem. You may contact the George Mason University Office of Research Integrity and Assurance at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research.

This research has been reviewed according to George Mason University procedures governing your participation in this research.
CONSENT
I have read this form and agree to participate in this study.

_____   _____   _____
Name (Print)   Name (Sign)   Date
Version date: January 13, 2015

2. INFORMED CONSENT FORM (FOR 18 +STUDENTS)

"**Understanding the Variable Effectiveness of Written Corrective Feedback in the**

Acquisition of Grammatical Forms: A Linguistic Perspective"

RESEARCH PROCEDURES
Mr. Wagner wants to do a study to see if the way he teaches writing helps students to become better writers. If you agree to participate, you will write 4 paragraphs. You must write these 4 paragraphs for class even if you do not participate in the study. If you say, "**Yes**," your paragraphs will be used in Mr. Wagner's study. If you say, "**No**," your paragraphs will <u>not</u> be used for Mr. Wagner's study.
RISKS
There are no foreseeable risks for participating in this research.
BENEFITS:
There will be no direct benefits to participants.
Mr. Wagner's hopes his research will be used to make other SLL writing classes better so that other SLL students in other SLL classes can learn to be better writers faster.
CONFIDENTIALITY
Your name will not be used in this study. The data in this study will be confidential. Names and other identifiers will be blacked out on writing samples: (1) your name will be blacked out on all the collected data; (2) a code will be placed the collected data; (3) through the use of an identification key, the researcher will be able to link your writing samples to your identity; and (4) only the researcher will have access to the identification key. Where names are written on writing samples, they will later be blacked out and will only be identified by participants' ID numbers.
PARTICIPATION
You do not have to participate in this study if you do not want to participate. You can just tell Mr. Wagner, "I do not want to be in the study." If you decide to participate now, but later decide that you do not want to be in the study, you can tell Mr. Wagner, "I do not want to be in the study anymore." This will not affect your grade in Mr. Wagner's class. You can change your mind at any time. You will not be penalized for not participating or for ending your participation.
CONTACT
This research is being conducted by Doug Wulf and Jason Wagner at George Mason University. You can call Doug at (703) 993-2771 and Jason at (703) 674-7194 for questions or to report a research-related problem. You may contact the George Mason University Office of Research Integrity and Assurance at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research.

This research has been reviewed according to George Mason University procedures governing your participation in this research.

CONSENT

I have read this form and agree to participate in this study.

_____   _____   _____

Name (Print)   Name (Sign)   Date

Version date: January 13, 2015

3. INFORMED CONSENT FORM (FOR PARENTS)

"Understanding the Variable Effectiveness of Written Corrective Feedback in the

Acquisition of Grammatical Forms: A Linguistic Perspective"

RESEARCH PROCEDURES
Mr. Wagner wants to do a study to see if the way he teaches writing helps students to become better writers. If you let your child participate, your child will write 4 paragraphs. Your child must write these 4 paragraphs for class even if he or she does not participate in the study. If you say, "**Yes**," your child's paragraphs will be used in Mr. Wagner's study. If you say, "**No**," your child's paragraphs will <u>not</u> be used for Mr. Wagner's study.

RISKS
There are no foreseeable risks for participating in this research.

BENEFITS
There will be no direct benefits to participants.
Mr. Wagner's hopes his research will be used to make other Spanish language learners (SLL) writing classes better so that other SLL students in other SLL classes can learn to be better writers in their second language faster.

CONFIDENTIALITY
Your child's name will not be used in this study. The data in this study will be confidential. Names and other identifiers will be blacked out on writing samples: (1) your child's name will be blacked out on all collected data; (2) a code will be placed on the collected data; (3) through the use of an identification key, the researcher will be able to link your child's writing samples to his or her identity; and (4) only the researcher will have access to the identification key. Where names are written on writing samples, they will later be blacked out and will only be identified by participants' ID numbers.

PARTICIPATION
Your child does not have to participate in this study. Participation is completely voluntary. If you sign and return this form to Mr. Wagner and your child also agrees to participate; your child's essays will be used in the study. If you do <u>not</u> sign and return this form, his or her essays will <u>not</u> be used in this study. You can contact Mr. Wagner at any time to end your child's participation in the study. There is no penalty for not participating or for ending participation.

CONTACT
This research is being conducted by Doug Wulf and Jason Wagner at George Mason University. You can call Doug at (703) 993-2771 and Jason at (703) 674-7194 for questions or to report a research-related problem. You may contact the George Mason University Office of Research Integrity and Assurance at 703-993-4121 if you have questions or comments regarding your child's rights as a participant in the research.

This research has been reviewed according to George Mason University procedures governing your participation in this research.

CONSENT

I have read this form and agree to participate in this study.

_____

Student's Name (Print)

_____   _____   _____

Parent's Name (Print)   Parent's Name (Sign)   Date

_____   _____   _____

Parent's Name (Print)   Parent's Name (Sign)   Date

Version date: January 13, 2015

E. **Student Writing Sample Sheet (Blank)**

Student Writing Sample

Class #_____

Writing Sample #_____

Student Name: _____ **Student #** _____ **Date:** _____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

REFERENCES

Afitska, O. (2015). Role of focus-on-form insruction, corrective feedback and uptake in second language classrooms: Some insights from recent second language acquistion research. *The Language Learning Journal, 43* (1), 57-73.

Ammar, A., & Lightbown, P. M. (2005). Teaching marked linguistic structure- More about acquisition of relative clauses by Arab learners of English. In A. Housen, & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 167-198). New York, NY: Mouton de Gruyter.

Anderson, J. (1976). *Language, memory, and thought.* New York: Lawrence Erlbaum Associates.

Anderson, J. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Anderson, J. (1990). *Cognitive psychology and its implications.* New York, NY: Freeman.

Anderson, J. (2005). *Cognitive psychology and its implications.* New York, NY: Worth.

Anderson, R. (1984). The one to one prinicple of interlanguage construction. *Language Learning , 34* (4), 77-95.

Anderson, R., & Shirai, Y. (1996). The primacy of aspect in first and second language

    acquisition: The Pidgin/Creole connection. In W. Ritchie, & T. Bhatia (Eds.),

    Handbook on language acquisition (pp. 527-571). New York, NY: Academic

    Press.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft

    composition classroom: Is the content feedback followed by form feedback the

    best method? *Journal of Second Language Writing, 9* (3), 227-258.

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form,*

    meaning and use. Malden, MA: Blackwell.

Bialystok, E., & Ryan, E. (1985). A metacognitive framework for the development of

    first and second language skills. In D. Forrest-Pressley, C. MacKinnon, & T.

    Waller (Eds.), *Metacognition, cognition, and human performance* (Vol. 1, pp.

    207-252). Orlando, FL: Academic Press.

Bird, S. (2010). Effects of distributed practice on the acquisition of second language

    English syntax. *Applied Psycholinguistics, 31* (4), 635-650.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of*

    Second Language Writing, 17 (2), 102-118.

Bitchener, J., & Knoch, U. (2008). The value of written corrective feeback for migrant

    and international students. *Language Teaching Research, 12*, 409-431.

Bitchener, J., & Knoch, U. (2009). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics, 31* (2), 193-214.

Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing, 19* (4), 207-217.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing, 14* (3), 191-205.

Bruton, A. (2010). Another reply to Truscott on error correction: Improved situated designs over statistics. *System, 38* (3), 491-498.

Butler, Y. (2002). Second language learners' theories on the use of English articles: An analysis of the metalinguistic knowledge used by Japanese students in acquiring the English aritlce system. *SSLA, 24* (3), 451-480.

Chandler, J. (2003). The efficacy of various kinds of error feedack for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing, 12* (3), 267-296.

Clark, H. (1977). Bridging. In P. Johnson-Laird, & P. C. Wason, *Thinking: Readings in cognitive science* (pp. 411-420). Cambridge: Cambridge University Press.

Collins, L., Trofimovich, P., White, J., Cardoso, W., & Horst, M. (2009). Some input on the easy/difficult grammar question. *Modern Language Journal, 93* (3), 336-353.

DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19* (2), 195-221.

DeKeyser, R. M. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing a second language. In C. Doughty, & J. Williams, *Focus on form in classroom second language acquistion* (pp. 42-63). New York, NY: Cambridge University Press.

DeKeyser, R. M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 125-51). New York: Cambridge University Press.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning, 55* (1), 1-25.

DeKeyser, R. M. (2007). Introduction: Situating the concept of practice. In R. M. DeKeyser (Ed.), *Practice in Second Language: Perspectives from applied linguistics and cognitive psychology* (pp. 1-18). Cambridge: Cambridge University Press.

DeKeyser, R. M., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics, 31* (3), 413-438.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences , 22*, 735-808.

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an emperical study of SL relativization. *Studies in Second Language Acquisition , 13* (4), 431-469.

Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197-262). Cambridge, England: Cambridge University Press.

Dulay, H., & Burt, M. (1978). Some remarks on creativity in language acquisition. In W. C. Richie (Ed.), *Second language acquisition research: Issues and impliciations* (pp. 65-89). New York: Academic Press.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language two.* Oxford, England: Oxford University Press.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27* (2), 315-330.

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

Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27*, 305-352.

Ellis, R. (1990). *Instructed second language acquisition.* Oxford: Blackwell.

Ellis, R. (1994). A theory of instructed second language acquisition . In *Implicit and explicit learning in language* (pp. 79-114). London, England: Academic Press.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27* (2), 141-172.

Ellis, R. (2009a). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, E. C., R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-25). Tonawand, NY: Multilingual Matters.

Ellis, R. (2009b). *The study of second language acquisition.* Oxford, England: Oxford University Press.

Ellis, R., Loewen, S., & Basturkmen, H. (2006). Disentangling focus on form. A response to Sheen and O'Neill (2005). *Applied Linguistics, 27* (1), 135-141.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System, 36* (3), 353-371.

Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In *Second language writing: Research insights for the classroom* (pp. 178-190). Cambridge: Cambridge University Press.

Ferris, D. (1999). The case for grammar correction in L2 writing classes: A reponse to
Truscott (1996). *Journal of Second Language Writing, 8* (1), 1-11.

Ferris, D. (2004). The "grammar correction" debate in L2 writing: Where are we, and
where do we go from here? (and what do we do in the meantime...?). *Journal of
Second Language Writing, 13* (1), 49-62.

Ferris, D. (2006). Does error feedback help student writers/new evidence on the short-
and long-term effects of written error correction. In K. Hyland, & F. Hyland,
Feedbck in Second Language Writing: Contexts and Issues (pp. 81-104).
Cambridge: Cambridge University Press.

Ferris, D. (2010). Second language writing research and written corrective feedback in
SLA. *Studies in Second Language Acquisition, 32* (2), 181-201.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does
it need to be? *Journal of Second Language Acquisition, 10* (3), 161-184.

Ferris, D., Chaney, S. J., Komura, K., Roberts, B. J., & Mckee, S. (2000). Perspectives,
problems, and practices in treating written error. *International TESOL
Convention.* Vancouver, B.C.

Gass, S., & Mackey, A. (2002). Frequency effects and second language acquisition: A
complex picture? *Studies in Second Language Acquisition, 24* (2), 249-260.

Gass, S., & Selinker, L. (1992). *Language transfer in language learning.* Amsterdam:
John Benjamins.

Givon, T. (1991). Markedness and grammar: Distributional, communicative and

    cognitive correlates of syntactic structure. *Studies in Language, 15*, 335-370.

Givon, T. (1995). *Functionalism and grammar.* Amsterdam: John Benjamins.

Goldschneider, J. M., & DeKeyser, R. M. (2005). Explaining the "natural order of L2

    morpheme acquisition" in English: A meta-analysis of multiple determinants.

    Language Learning, 55 (1), 27-77.

Guenette, D. (2007). Is feedback pedagogically correct?: Research design issues in

    studies of feedback on writing. *Journal of Second Language Writing, 16* (1), 40-

    53.

Han, Z. (2013). Forty years later: Updating the fossilzation hypothesis. *Langauge

    Teaching, 46* (2), 133-171.

Harrington, M. (2004). Commentary: Input processing as a theory of processing input. In

    B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary*

    (pp. 79-92). Mahwah, NJ.

Hinkel, E. (2002). Why English passive is difficult to teach (and learn). In E. Hinkel, &

    S. Fotos (Eds.), *New perspectives on grammar teaching in second language

    classrooms* (pp. 233-259). Mahwah, NJ: Lawrence Erlbaum.

Housen, A., Pierrard, M., & Van Daele, S. (2005). Structure complexity and the efficacy

    of explicit grammar instruction. In A. Housen, & M. Pierrard (Eds.),

Investigations in instructed second language acquisition (pp. 235-269). Brusells, Belguim: Mouton De Gruyter.

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research , 18*, 193-223.

Hulstijn, J., & de Graaff, R. (1994). Under what conditions does explicit knolwedge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review, 11*, 97-112.

Izquierdo, J., & Collins, L. (2008). The facilitative role of L1 influence in tense-aspect marking: A comparison of Hispanophone and Anglophone learners of French. *Modern Language Journal, 92*, 350-368.

Izumi, Y., & Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English. *The Canadian Modern Language Review, 60* (5), 587-609.

Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8* (1), 63-99.

Kepner, C. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *Modern Language Journal, 75* (3), 305-313.

Kharma, N. (1981). Analysis of the errors committed by Arab university students in the use of the English definite/indefinite articles. *IRAL, 19* (1), 333-345.

Kim, I., LaPointe, L. L., & Stierwalt, J. A. (2012). The effect of feedback and practice on

  the acquisition of novel speech behaviors. *American Journal of Speech-Language*

  Pathology, 21 (2), 89-100.

Krashen, S. D. (1981). *Second language acquistion and second language learning.*

  London, England: Pergamon.

Krashen, S. D. (1982). *Principles and practice in second language acquisition.*

  Englewood Cliffs, NJ: Prentice-Hall.

Krashen, S. D., & Terrell, T. (1983). *The natural approach.* New York, NY: Pergamon.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.*

  Ann Arbor, MI: University of Michigan.

Lalande, J. (1982). Reducing composition errors: An experiment. *Modern Language*

  Journal, 66, 140-149.

Lightbown, P. (1992). Getting quality input in the second/foreign langauge classroom. In

  C. Kramsch, & S. McConnell-Ginet, *Text and context: Cross-disciplinary*

  perspectives on language study (pp. 187-97). Lexington, MA: D.C. Health.

Lightbown, P. (1998). The importance of timing in focus on form. In C. Doughty, & W.

  Jessica, *Focus on form in classroom second language acquistion* (pp. 177-96).

  Cambridge, England: Cambridge University Press.

Lightbown, P. M., & Spada, N. (2000). Do they know what they're doing? L2 learners'

  awareness of L1 influence. *Language Awareness , 9* (4), 198-216.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition, 27* (3), 361-386.

Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21* (4), 557-587.

Mackey, A., & Philip, J. (2011). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal, 82* (3), 338-356.

Master, P. (1991). Active verbs with inanimate subjects in scientific prose. *English for Specific Purposes , 10*, 15-33.

Master, P. (1994). The effect of systematic instruction on learning the English article system. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 229-252). New York, NY: Cambridge University Press.

McLaughlin, B. (1987). *Theories of second language acquistion.* London: Edward Arnold.

Meisel, J., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, 3* (2), 109-135.

Nagata, N. (1998). The relative effectiveness of production and comprehension practice

    in second language acquistion. *Computer Assisted Language Learning, 11* (2),

    153-177.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and

    quantitative meta-analysis. *Language Learning, 50* (3), 417-528.

Odlin, T. (1989). *Language transfer.* Cambridge, England: Cambridge University Press.

Odlin, T. (2003). Cross-linguistic influence. In C. Doughty, & M. Long (Eds.), *The*

    handbook of second language acquisition (pp. 436-486). New York, NY:

    Blackwell

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications

    for bilingualism and SLA. In *Implicit and explicit learning of languages* (pp. 393-

    419). San Diego, CA: Academic Press.

Paradis, M. (2009). *Declarative and procedural determinants of second languages.*

    Amsterdam: John Benjamins.

Pienemann, M. (1989). "Is language teachable?" Psycholinguistic experiments and

    hypotheses. *Applied Linguistics, 10* (1), 52-79.

Pienemann, M., Johnson, M., & Brindley, G. (1988). Constructing an acquisition-based

    procedure for assessing second language acquisition. *Studies in Second Language*

    Acquisition, 10 (2), 217-243.

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing, 21* (4), 375-389.

Polio, C., Fleck, C., & Leder, N. (1998). 'If only I had more time': ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing, 7* (1), 43-68.

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly, 20*, 83-93.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18* (1), 27-68.

Rogers, D. (2011). The automaztization of verbal morphology in instructed second language acquisition. *IRAL, 49* (4), 295-319.

Scheffler, P. (2009). Rule difficulty and the usefulness of instruction. *ELT Journal, 63* (1), 5-12.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge, England: Cambridge University Press.

Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition , 15*, 147-163.

Semke, H. (1984). The effects of the red pen. *Foreign Language Annals, 17* (3), 195-202.

Sharwood-Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics , 2*, 159-168.

Sheen, R. (2003). Focus on form- A myth in the making? *ELT Journal, 57* (3), 225-233.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TELOL Quarterly, 41* (2), 255-83.

Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition, 32* (2), 203-234.

Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and metalinguisitic explanation on learners' explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing, 22* (3), 286-306.

Shirai, Y. (2004). A multi-factor account for form-meaning connections in the acquisition of tense-aspect morphology. In B. VanPatten, S. Williams, & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp. 97-122). Mahwah, NJ: Lawrence Erlbaum.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford, England: Oxford University Press.

Smith, M. (1993). Input enhancement in instructed SLA: Theoretical basis. *Studies in Second Language Acquisition, 15* (2), 165-79.

Spada, N., & Lightbown, P. M. (1993). Instruction and the development of questions in the L2 classroom. *Studies in Second Language Acquisition, 15* (2), 205-221.

Spada, N., & Lightbown, P. M. (1999). Instruction, L1 influence and developmental readiness in second language acquisition. *Modern Language Journal , 83*, 1-22.

Spada, N., & Lightbown, P. M. (2008). Form-focused instruction: Isolated or integratd? *TESOL Quarterly , 42*, 181-207.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60* (2), 263-308.

Stockwell, R. P., Bowen, J. D., & Martin, J. W. (1965). *The grammatical structures of English and Spanish.* Chicago, IL: University of Chicago Press.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning, 46* (2), 327-369.

Truscott, J. (1999). The case for "the case for grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing, 8* (2), 111-122.

Truscott, J. (2004). Evidence and conjecture: A response to Chandler. *Journal of Second Language Writing, 13* (4), 337-343.

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing, 16* (4), 255-272.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition.* Norwood, NJ: Greenwood Publishing Group.

VanPatten, B. (2002). Processing instruction: An update. *Language Learning, 52* (4), 755-803.

VanPatten, B. (2004). Input processing in second language acquisition. *Processing Instruction: Theory, Research, and Commentary, 9* (3), 5-31.

VanPatten, B. (2007). Input processing and adult SLA. In *Theories of second language acquisition: An introduction*. In B. VanPatten & J. Williams (Eds.), *Theories of Second language acquisition* (pp. 115-135). Mahwah, NJ: Lawrence Erlbaum.

Wagner, J. (2013, May). Defining grammatical difficulty in the L2-writing classroom: Examining students' grammatical accuracy improvements to establish a measure of grammatical difficulty. *Unpublished Qualifying Paper (George Mason University)* . Fairfax, VA.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research , 7*, 133-161.

White, L. (2003). Fossilization in steady state L2 grammars: persistent problems with inflectional morpology. *Bilingualism: Language and Cognition, 6* (2), 129-141.

White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness , 11* (4), 259-290

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing, 21* (4), 321-331.

Yamada, J., & Matsuura , N. (1982). The use of the English article among Japanese

 students. *RELC Journal*, *13* (1), 50-63.

Yip, V. (1995). *Interlanguage and learnability: From Chinese to English.* Amsterdam:

 John Benjamins.

Zephir, F. (2000). Focus on form and meaning: Perspectives of developing teachers and

 action-based research. *Foreign Language Annals, 33* (1), 19-30.

BIOGRAPHY

Jason Wagner grew up in Windber, Pennsylvania, a small town just east of Pittsburgh. He graduated from Johnstown Christian School in 1994. He received his Bachelor of Arts from Campbell University in 1998. He has taught or administered language programs in Taiwan, Europe, and United States for over 14 years. He received his Master of Arts in English/Linguistics from George Mason University in 2009. He recently moved to Kansas where he continues to teach English language learners.