STARS

2015

Increasing Metalinguistics Awareness as a Necessary Precursor for Preservice Teachers

Aimee Schoonmaker

University of Central Florida

University of Central Florida

STARS
Showcase of Text, Archives, Research & Scholarship

INCREASING METALINGUISTIC AWARENESS AS A NECESSARY PRECURSOR FOR
PRESERVICE TEACHERS

by

AIMEE C. SCHOONMAKER
B.A. Universidade Federal de Minas Gerais, 2005
M.A. University of Central Florida, 2010

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the College of Education and Human Performance
at the University of Central Florida
Orlando, Florida

Fall Term
2015

Major Professor: Kerry L. Purmensky

ABSTRACT

Metalinguistic awareness (MA) in the context of the present study is the ability to deconstruct and analyze a language's intricate systems, and by doing so, better understand how these systems work. More specifically for the present study, the investigation focused on preservice teachers' MA in relation to grammar and its importance for their future students' understanding of language. Language is at the core of any content area in students' academic lives and it will continue to permeate all aspects of their studies at all ages. As such, language-related issues should be at the front and center of preservice teachers' preparation.

The present study, a quasi-experimental one-group pretest-posttest (Shadish, Cook & Campbell, 2002) investigated the metalinguistic awareness of preservice teachers who were enrolled in a face-to-face, undergraduate applied linguistics course at an urban research university in the United States. The metalinguistic awareness in the scope of the present study was directly related to the grammar knowledge of participants, and it was measured by an adapted instrument called ALAT.

The results of the current study demonstrate that *overall increase* of MA is feasible (as attested by the results from research question 1). Nevertheless, further investigations (research questions 2, 3, 4, and 5) demonstrated that teachers' levels of metalinguistic awareness vary significantly. These results echoed previous findings that demonstrated that preservice teachers are not language-knowledgeable enough to deal with myriad issues that involve language, both in relation to students who are native speakers of English and also concerning ELs (Kolln & Hancock, 2005; Nutta et al., 2012; Pappamihiell, 2007). Preservice teachers' lack higher level

metalinguistic awareness, as evidenced by their limited ability to explain grammar errors and use proper metalanguage while doing so.

The main implication of the present study lies in the recommendation that more can and should be done in order to ensure that preservice teachers are receiving the appropriate amount of language-oriented preparation during their college years. The present study offers a confirmatory perspective to previous research findings which found that preservice teachers are not knowledgeable enough in relation to language. Previous studies also pointed out to this lack of preparation as a generator of feelings of inadequacy and anxiety in preservice teachers while foreseeing their future students' language struggles. Nonetheless, the present study also demonstrates that improvement can be achieved in regard to MA teacher preparedness in relatively little amount of time, especially for recognition of grammatical items such as parts of speech and parts of sentence. However, the findings demonstrate that more time is needed to ensure better results for complex grammar analyses such as explanation of grammatical errors.

I dedicate this journey to my lovely kids Carol and Sam, and to my husband Tracey.

Each one, in his or her way, journeyed with me.

ACKNOWLEDGMENTS

A Ph.D. journey can be simultaneously thrilling and daunting. However, it is never one without many lessons learned, and in my case, it certainly required large amounts of perseverance and much appreciated support from many individuals. The lessons learned were priceless, and I am sure I will be able to apply them to many other areas of my life.

I would like to thank Dr. Kerry Purmensky, my chair, who was also my mentor for many years while I was a Graduate Teaching Assistant at the University of Central Florida. In this position, she never failed to supply me with indispensable advice and provide me with a role model for ethical and responsible conduct. I thank her for her unrelenting support, not only as the Chair of my dissertation, but also for her guidance in my professional development. She has been a powerful drive propelling me to take new challenges and learn new paths and to do so without the fear of the unknown. "Dr. P." is contagiously optimistic, and that quality certainly helped me in difficult times. Her focus and objectivity, both qualities that I am not famous for having, also made sure I would divert less while working through my research.

My committee members were chosen carefully, and I am so thankful for the fact that not only did they accept to be on my committee, but they also made themselves available and generously provided advice when needed. Dr. Florin Mihai was my direct supervisor for five years while I taught TSL4240, Issues in Second Language Acquisition, at the University of Central Florida. In this role, he gave me enormous freedom to develop my skills as an instructor, and his confidence in my competence instilled me with a great sense of independence. I thank him for being always such a positive and solid presence throughout my academic and

professional years at UCF and for his objective advice on crucial aspects of my dissertation. Dr. Joyce Nutta has been a source of inspiration to all of us in the first TESOL Ph.D. cohort at the University of Central Florida. From her passionate stance on helping ELLs develop to the best of their abilities to her tireless commitment to the TESOL cause, Dr. Nutta has been showing us, with her own example, that we can thrive as successful professionals and spread the word out as fully-fledged TESOLers. Her enthusiasm for each one of us who got to the end of the line in this journey is genuine and heartfelt. Dr. Keith Folse has been a constant driving force to all of us who were lucky to be one of his students, and who understood the importance of grammar through *his* mastery of grammar, and how to approach grammar from an engaging and effective perspective. I thank him for his time and expertise in the early stages of my dissertation process, when his experience guided some of my decisions in how to pursue my topic. Dr. M. H. Clark's attention to detail and her ability to dissect complicated analyses in ways that are accessible and clear were crucial to my progress, first as her student, and later as her mentee for my dissertation statistical analyses. I am so very thankful for her expertise and availability during my dissertation, when she generously dedicated time, energy, and patience to answer my questions, discuss my analyses, and suggest new ways of exploring my research.

On a personal level, I would like to acknowledge Dr. Marcella Farina, who became a dear friend and mentor since the very beginning of our Ph.D. journey. It has been a privilege to have access to her many areas of expertise and to have her constant support. Her guidance was also fundamental for the interrater reliability and rating decisions for this study, as well as her role as a second rater. I thank her for her time, her patience, and her constant support. I would not be here without her, and I hope that one day I will be able to be a mentor to someone in the same

Finally, I thank my family, which is always at the front and center of my life. There is no victory for me that is not ours. I thank my husband Tracey, who never failed to instill in me the need to keep moving to the next dream and to the next challenge, and do so with energy and confidence. His discipline and drive have always offered me inspiration to pursue my goals. I thank my children, Caroline and Samuel, because they have always been my number one drive in life. If my personal journeys can inspire them in any possible way to also pursue their dreams, then everything was worth it. I sincerely thank them for their love and support, and for their unrelenting belief that I can continue to be their (little) hero.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS & ACRONYMS

EFL	English as a Foreign Language
EL	English Learner
ELL	English Language Learner
ESL	English as a Second Language
ESOL	English for Speakers of Other Languages
GTA	Graduate Teaching Associate
KaL	Knowledge about Language
L1	First, or Native, Language
L2	Second, or Target, Language
LA	Language Awareness
MA	Metalinguistic Awareness
SLA	Second Language Acquisition
TEFL	Teaching English as a Foreign Language
TESL	Teaching English as a Second Language
TMA	Teacher Metalinguistic Awareness
TESOL	Teaching English to Speakers of Other Languages

CHAPTER 1:

INTRODUCTION TO THE STUDY

Language can be considered one of the most unique characteristics of the human species, and although languages are shaped in symbolic and quite arbitrary systems, most humans make the transition from no language stages to fully-fledged conversationalists in quite an effortless mode, usually without any *need* for formal guidance or conscious practice (Chomsky, 1968).

This does not mean that children receive no support at all throughout this language acquisition process. Infant-caregiver interaction provides both critical input and feedback that are at the basis of socialization, and that will help shape children's progress during their initial developmental processes (Skinner, 1957). Nonetheless, guidance (alone) would not justify their language acquisition being so fluid and without major roadblocks. Human beings *acquire* their first languages, rather than *learn* them, and are fully equipped to do so in an innate fashion, granted that no other physical or neurological constraints might impair this process to occur (Chomsky, 1968).

However, if language acquisition to prime functional levels seems to be a given for most adult language users, the same is not true in relation to a better structural knowledge and deeper understanding of language forms. The mastering of a language's sophisticated rule-oriented systems is spontaneous, without conscious intention. It is also effective for one's own production of language as well as the understanding of language produced by others (Chomsky 1968). However, the same cannot be said about the ability to *deconstruct* language, analyze it, and observe it critically.

Metalinguistic Awareness

Metalinguistic awareness (MA) can be defined as the ability to understand language in a more explicit fashion, in a way being able to objectify and manipulate it beyond the levels of pure functional use. Such ability exceeds mere knowledge *about* language, but it also encompasses both sensitivity and consciousness about the *nature* of language (Donmall, 1985). Moreover, and crucial to the present study, is the fact that teachers' MA is an essential component of their professional skills, and should be also paramount within the preservice preparation scope (Moats, 1994, 2009; Kolln & Hancock, 2005; Myhill, 2011; Myhill, Jones, & Watson, 2013).

The notion of a specific kind of language awareness has been the object of scrutiny in different areas of study for decades, among them the fields of applied linguistics, cognitive science, cognitive psychology, and education. However, a precise definition of such awareness, or one that is shared and agreed upon by all fields, has never been easily reached. In spite of great interest in the matter, linguists, cognitive scientists, psychologists, and education professionals, for example, have hardly agreed upon one absolute definition of MA, which is in itself a testament to its complexity.

Multiple fields conceptualized this awareness about language with different nuances, but the term *metalinguistic awareness* has been consistently favored, and achieved a solid status as a sound definition of what awareness about language encompasses. MA has been theoretically framed under a variety of perspectives, in direct relation but not restricted to, grammar, phonology, phonetics, syntax, morphology, and bilingualism, just to name a few topics (Andrews, 1997, 1999a, 1999b, 2001, 2006; Apel, Fowler, Brimo, & Perrin, 2012; Bialystok &

Barac, 2012; Ellis, 1991, 2004; Hun & Ellis, 1998; Moats, 1994; Roehr & Ganem-Gutierrez, 2009).

MA, in the context of the present study, espouses Myhill's views of language being treated as an artifact and the object of "conscious monitoring and manipulation" (Myhill, 2011, p. 250). It also endorses the idea of *grammatical content* knowledge as belonging to an overarching notion of linguistic knowledge, or metalinguistic content knowledge. A more specific aspect of relevance is the notion of an explicit knowledge of grammar, specifically in relation to morphology and syntax (Myhill, 2013).

MA for Preservice Teachers. The need for an enhanced awareness about language with is an essential precursor in the academic preparation stages for preservice teachers, whose need for a more robust understanding of language structures is much more compelling than that of the laymen. In a way, such a perspective also invites a look back at children's first literacy developmental stages, when metalinguistic abilities play a salient role (Montrul, 2008; Tunmer, 1984; Zipke, 2008, 2011).

Preservice teachers' preparation years should be a prime time to foster language-confident professionals, who could in turn face their future students' language needs with enthusiasm, not fear or anxiety (Pappamihiell, 2007).Their preparation should offer plenty of opportunities for language training exposure and consequent language improvement, but that is not always the case, as research findings have demonstrated (Kolln & Hancock, 2005). It is in this vein that the present study proposes to investigate MA with specific focus on grammar as a part of a preservice teacher preparation program.

Increasing MA is not only beneficial but an essential step for future educators dealing with language (Andrews, 1997, 1999a, 1999b, 2001, 2003; Moats, 1994, 2009; Myhil et al., 2013).While investigating MA in the specific sphere of English language teachers, research findings have also found that teachers' metalinguistic knowledge of English is a critical component of their teaching effectiveness in the language and that *lack* of MA may in fact compromise students' success rates (Andrews, 1997, 2001, 2003; Moats, 1994, 2009; Myhill et al., 2013; Wright 1991, 2002). As a result, this knowledge base is an indicator of how preservice teachers will be able to meet the strenuous standards in the language classroom (Nutta et al., 2012).

Are We Preparing Metalinguistically-Aware Teachers? Research findings suggest that much has to be improved about how teacher preparation courses allot time for solid and specific language training for teachers (Andrews, 1997, 1999a, 1999b, 2001, 2003; Moats, 1994, 2009; Myhil et al., 2013). Oftentimes, teacher preparation in general sadly dedicates "more attention to aspects of methodology and class management than to language *per se"* (Andrews, 2003, p. 82). Why does that happen? Could the reason be that institutions tend to rely on an assumption of proficiency on the part of the teachers, especially the native speakers? Is there a perception in the teacher preparation realm that native speakers are competent to teach language solely for being native speakers?

There seems to be more attention directed to *students*' struggles than has ever been paid to *teacher* preparation concerning language knowledge. In spite of that, teachers have expressed clear insecurities about their preparedness to teach language-oriented topics, especially to ELs (Nutta et al., 2012). Research on how preservice teachers feel about teaching ELs demonstrates

that preservice teachers expect to face more language issues with English learners than with native speaker students in their classroom s(Pappamihiel, 2007).

In a thorough article analyzing the story of English grammar in the United States, Kolln and Hancock (2005) go to great lengths to depict the endless discussion, debate, disagreement and ambivalent views, both in favor and against grammar teaching in the country. They also recognize how monumental the task of achieving consensus on the need for grammar teaching is and how difficult it is to define the format in which it should be implemented throughout the country. Notwithstanding, they strongly condemn the "appalling lack of training for teachers and prospective teachers" in this critical area (p. 29).

A Balanced View on Preservice Teacher Preparation. Nutta et al. (2012) believe that in relation to English Learners (ELs) and their teachers, there is a need for augmented teacher preparation that should include both an understanding of the nature of language, as well as how "to support second language development and academic achievement" (Nutta et al., 2012, p. 1). In this same vein, and crucial to the present investigation, is the notion that although those requirements are made in relation specifically to ELs, they are as valid for any student in mainstream classes, including native speakers of English. By the same token, the need for *teachers* to be well prepared about language issues should not be limited to the ones working with ELs, but be recognized as relevant to any teacher who will be in charge of students at various stages of literacy development (Moats, 1994; Tunmer et al., 1984).

If previous methods for teaching grammar have consistently failed to show positive results (Ellis, 2006), that should not suggest that all investigations of patterns of language should be shunned in the linguistic field (Kolln & Hancock, 2005). Bloor (1986) insists on a more

5

positive view of grammar instruction: "…formal grammar is an important aspect, and with all its faults, which are grievous, traditional grammar is better than no grammar at all" (p. 158). The same notion was espoused by Crystal (2006), who not only points to the relevance of grammar, but makes a case for an approach that can be engaging and positively challenging for students. That approach is valid and true to a number of possible areas, not exclusively for those pursuing specifically language-driven professions but for any educator who will be in charge of a classroom with learners of any age and language background.

The need to assist preservice teachers in order to develop their own language skills draws from the same vein of what will be expected of their future students. In what is being considered a trend for the ESL classroom, content-based instruction has pointed to a clear change from teaching "language for language's sake" to an approach that embraces grammar (and vocabulary) as essential in achieving the necessary language arts standards at grade-level proficiency (Nutta et al., 2012, p. 9).

Henceforth, it is argued for the purpose of this investigation that preservice teachers are not receiving enough language exposure to develop MA. As future educators in charge of diverse classrooms, it is not realistic to assume that preservice teachers themselves have optimal MA, something that has been clearly demonstrated in research findings. Lack of MA, in the present study with a focus on grammar, is a problem that can be minimized through an enhanced language-focused curriculum for preservice teachers, and its implementation should be based on empirical and solid research.

In the present study, participants were preservice teachers who were tested with an instrument called Adapted Language Awareness Test (ALAT). The ALAT progresses from mere

recognition of grammatical items to deeper levels of language production, investigating

participants' performances from pre- to posttest, and intermediated by a treatment in the form of

a three-week grammar module. The first research question investigated the overall ALAT results,

and the subsequent research questions investigated each one of its parts.

Research Questions

1. Is there a statistically significance increase in the overall MA of participants after a treatment, as measured by pre- and post-administrations of the ALAT?

2. Is there a statistically significant difference in the ability of participants to *identify parts of speech* after a treatment as measured by pre- and post-administrations of the first part of the ALAT?

3. Is there a statistically significant difference in the ability of participants to *identify parts of sentence* after a treatment as measured by pre- and post-administrations of the second part of the ALAT?

4. Is there a statistically significant difference in the ability of participants to *identify* and *correct grammatical errors* in a sentence after a treatment as measured by pre- and post-administrations of the third part of the ALAT?

5. Is there a statistically significant difference in the ability of participants to *provide* metalinguistic *explanations for grammatical errors* after a treatment as measured by pre- and post-administrations of the fourth part of the ALAT?

Research Hypotheses

1. On the subject of Research Question 1, the directional hypothesis states that there is a statistically significant increase ($p < .05$) in the overall metalinguistic awareness of

participants after a treatment, as measured by pre- and post-administrations of the ALAT. The null hypothesis (H0: $\mu1- \mu2= 0$) states that there is no difference between the scores of pretest and posttest of the ALAT after the treatment. The null hypothesis is expected to be rejected in favor of an overall increase in performance (Ha: $\mu1 - \mu2 > 0$).

2. On the subject of Research Questions 2, the non-directional hypothesis states that there is a statistically significant difference ($p < .05$) between the scores of a pre- and post-administration of the ALAT in relation to identification of grammatical items, mostly parts of speech. The null hypothesis (H0: $\mu1- \mu2= 0$) states that there is no difference between the scores of pretest and posttest of the ALAT after the treatment. The null hypothesis is expected to be rejected in favor of a difference in performance (Ha: $\mu1 - \mu2 \neq 0$).

3. On the subject of Research Question 3, the non-directional hypothesis states that there is a statistically significant difference ($p < .05$) between the scores of a pre- and post-administration of the ALAT in relation to identification of parts of sentence. The null hypothesis (Ha: $\mu1- \mu2= 0$) states that there is no difference between the scores of pre- and posttest of the ALAT. The null hypothesis is expected to be rejected in favor of a difference in performance (Ha: $\mu1 - \mu2 \neq 0$).

4. On the subject of Research Questions 4, the non-directional hypothesis states that there is a statistically significant difference ($p > .05$) between the scores of a pre- and post-administration of the ALAT in relation to ability to identify and correct grammatical mistakes. The null hypothesis (Ha: $\mu1- \mu2= 0$) states that there is no

difference between the scores of pre- and posttest of the ALAT. The null hypothesis is expected to be rejected in favor of a difference in performance (Ha: $\mu1 - \mu2 \neq 0$).

5. On the subject of Research Question 5, the non-directional hypothesis states that there is a statistically significant difference ($p > .05$) between the scores of a pre- and post-administration of the ALAT in relation to ability to name and explain the same aforementioned grammatical mistakes. The null hypothesis (Ha: $\mu1 - \mu2 = 0$) states that there is no difference between the scores of pre- and posttest of the ALAT. The null hypothesis is expected to be rejected in favor of a difference in performance (Ha: $\mu1 - \mu2 \neq 0$).

ALAT Overview. The ALAT is composed of four distinct parts. These four parts were organized under two distinct Sections, 1 and 2. Section 1 is oriented to recognition of grammatical terms, while Section 2 is dedicated to ungrammatical sentences. Each one of these parts was investigated by one of the Research Questions for the present study, and the results reported from pre- to posttest after a treatment. Research Question 1 investigated the *overall results*, encompassing all the combined results of the four parts of the ALAT. Research Question 2 investigated only Section 1, Part 1 of the ALAT, called S1P1. Research Question 3 investigated only Section 1, Part 2 of the ALAT, called S1P2. Research Question 4 investigated only Section 2, Part 1 of the ALAT, called S2P1. Finally, Research Question 5 investigated only Section 2, Part 2 of the ALAT, called S2P2.

In relation to its content, the ALAT starts from mere recognition of *parts of speech* (S1P1) and of *parts of sentence* (S1P2), to correction of *ungrammatical sentences* (S2P1) and culminates with specific explanations for the same *ungrammatical sentences* (S2P2). The first

and last parts of the ALAT vary enormously in difficulty, and such disparity can be better investigated by looking at the individual results for each section and task. Although difference in results was hypothesized from pre- to posttest after the treatment, it is understood that recognizing grammatical terms from a sentence differs dramatically from providing grammatical explanations for ungrammatical sentences. The issue of duration of the treatment and availability of enough opportunities to foster grammar reflection and expertise then becomes a source of questioning. Time could be insufficient to understand and explain more intricate aspects of grammar, such as different uses of personal pronouns (object and subject), or a subject-verb agreement error, just to cite a few examples present in the final part of the ALAT (S2P2).

The ALAT pretest version contained a bio-data questionnaire to inquire about items such as age, gender, major, nationality, and languages, as well as if participants had taken a specific course that features grammar training. Although these factors were not investigated in the statistical analyses nor were they part of the research questions, the purpose of this questionnaire is to narrow the scope of the study by providing a better picture of the demographics of the participants, consequently allowing for generalizability for a larger but similar population (Dörnyei, 2007).

Definition of Terms

It is essential to define the most relevant terms that are used throughout the study in order to better understand the proposed research questions and hypotheses. The main focus of the present study is the concept of *metalinguistic awareness* on grammar, as it relates to the competence demonstrated by participants, with specific focus on *preservice teachers*. The term

inservice will also be used sparingly in reference to research studies being depicted throughout the paper.

For the purpose of this study, MA is being conceptualized as one's ability to deconstruct language in relation to grammar, being able to identify parts of sentence and parts of speech, and not only recognize and identify errors but also correct them and provide accurate metalanguage (e.g. grammatical terminology) while explaining them (Ellis, 2006). The study investigated the current *metalinguistic awareness* of *preservice teachers* in an undergraduate course in relation to grammar knowledge on both ends of the spectrum of knowledge: the content-based and declarative knowledge as well as the procedural one, as proposed by Andrews (1997, 1999a, 1999b, 2001).

Another critical term present in this study was *participant*, mainly represented by a population of college undergraduates taking a specific course as part of their *preservice teacher* preparation. *Preservice teachers* who seek bachelor degrees in education from this university can pursue majors in early childhood education, elementary education, and English language arts, while also graduating with an ESOL endorsement. The aforementioned ESOL-endorsement course has a mostly applied linguistics-oriented content and is one of the mandatory steps of teacher preparation of *preservice teachers*. This course is designed to coveressential areas that equip students to teach English Learners (ELs).

These *preservice teachers* represented the main population for which the study aims at providing generalizability. Another subgroup of participants from a different academic population came from TEFL certificate courses, and they take this course as part of their pursuit of a certification in TEFL. A final subgroup is that of participants who take this applied

11

linguistics course as an elective course. In order to avoid misrepresenting these last two (admittedly less representative) groups also taking part in the study, the three groups combined will be referred to as *participants,* instead of *preservice teachers.* As stated before, the term *inservice* teachers may also be used throughout the paper in comparisons or in studies depicted in the paper, but not in relation to most of participants, who are *preservice*. A detailed description of these three groups is provided in the *Participants* section of the Methodology chapter.

A fundamental element in the study is the *Language Awareness Test*, which was used to measure how much MA participants have, as well as the ability to express such awareness. In the present study, the *Language Awareness Test* is operationalized as ALAT, an adaptation of a test first used by Bloor (1986), and subsequently by other researchers (Alderson, Steel, & Clapham, 1997; Andrews, 1999b, 2006, 2007; Alderson & Hudson, 2012). A detailed explanation and description of the ALAT as well as an overview of some of the previous studies that used the Language Awareness Test will be provided in the Literature Review and a summary in the *Instrumentation* section of the Methodology chapter.

The *ALAT,* as it is used in this study, consists of four parts, organized under two sections, Sections 1 and 2. Section 1 contains two separate parts. Part1: contains one sentence from which fourteen different grammatical categories and functions must be extracted, mostly parts of speech e.g. verb, noun, adjective, adverb, preposition, among others, and Part 2: contains four sentences from which participants must identify four specific parts of sentence: subject, predicate, direct and indirect object. Section 2 is also divided into two different parts: Part 1 contains 15 English sentences, each of them containing a mistake. For each sentence, participants were required to (a) underline the faulty part of the sentence; and (b) rewrite the faulty part of the

sentence correctly. Part 2 contains the same previous 15 sentences. For each sentence participants must provide an explanation for the rule that has been broken, ideally using adequate grammar terminology for maximum points. In order to make the references to these four parts more easily recognized, they were called throughout the study S1P1, S1P2, S2P1, and S2P2. Each of these parts corresponded to the four subsequent Research Questions 2, 3, 4, and 5. Research Question 1 was solely related to the combined *overall* results of the ALAT's four parts.

In the present study, the term *major areas of study* refers mostly, but not limited to the following three areas of study: (a) Early Childhood Development and Education; (b) Elementary Education; and (c) English language Arts Education. As stated before, less prevalent but also present as students for the same course are participants seeking a certificate in Teaching English as a Foreign Language (TEFL), which qualifies them to teach English as a foreign language both in the United States as well as overseas. As such, students from any other *majors* who are interested in teaching abroad may register for the certificate of which the applied linguistics course is one of the offered courses. A detailed account of the *Participants* for the present study will be provided in the Methodology chapter.

An initial portion of the ALAT instrument as a pretest contained a questionnaire to elicit bio-data relevant to participants' specific demographics, so that a better picture of the individual characteristics of the participants in the study was provided. In this questionnaire, participants detailed their majors, ages, gender and languages. The posttest did not have the bio-data portion.

Limitations of the Study

The first limitation of this study relates to the fact that what is being measured is admittedly a very small fraction of what language knowledge as a whole encompasses, and it

must be seen as mere sample of important notions that preservice teachers should be able to master before entering their future classrooms. The attempt to measure a complex construct such as *metalinguistic awareness* represents an ambitious, perhaps even impossible task. An accurate definition of this term may prove difficult. The amount of debate and disagreement concerning the scope of the language-grammar dyad bears witness to that. Also, the idea of *knowledge of language* seems to encompass the notion of *knowing* all there is to know about a language, which is an impossible goal to achieve and therefore cannot be measured. The grammatical items present in this study do not represent the entirety of what is believed to be the ideal body of language knowledge for professionals in teaching careers, but they represent a solid start.

Essentially, they do illustrate an essential portion of such knowledge, and a representative one based on K-12 standards for the state of Florida. As pointed out by Nutta et al. (2012), even without being ESL specialists, mainstream Language Arts and literacy teachers are bound to interact with ELs, not to mention the fact that for some schools the ideal scenario of having an ESL specialist is not a reality, which makes mainstream teachers the *only* teachers an English learner might have. This reality can only reinforce the need for preservice teachers to have a solid language knowledge base, and feel confident to guide their students in achieving the same.

A second limitation of this study concerns the number of participants available at this point and how generalizable the results can be, considering the wider population of preservice teachers across the state or even country. It is suggested as a possibility for future research that the same format be repeated in a series of studies for at least two years, instead of one term of classes, to enhance and solidify the methodology, as well as obtain a substantially larger sample size.

Potential Contribution of the Study

The main contribution of this study is to increase the empirical body of language research available to understand the current state of MA in the teacher preparation scope, bringing it to a positive and inclusive light. Instead of ignoring the issue of grammar language knowledge preparedness for preservice teachers, this investigation generates a useful approach to MA, pointing to its use as a tool to help teachers assist their future students in their language development, whether they are native or nonnative speakers of English.

Studies investigating MA for English as a foreign language environments tend to be more prevalent (Borg, 2003), but the same cannot be said about investigations of grammar in the teacher preparation field (Myhill et al., 2013). Under this premise, the pre- and post-ALAT could become a natural diagnostic step in the first or second year of the teacher preparation courses for all preservice teachers, ensuring that a solid language knowledge base is provided to all undergraduate students. In addition, the data for these years of testing could provide important information and impact how preservice teachers are being prepared for their future careers, not limited only to the issue of dealing with ELs, but also to the entire scope of their professional preparedness.

This study provides a wider and more in-depth body of research on how much opportunity preservice teachers have to increase their MA, and whether this amount of exposure is sufficient or not. The MA and language knowledge of preservice teachers should and can be improved, and this study examines this notion as a fundamental aspect through a structured and objective investigation.

CHAPTER 2:

LITERATURE REVIEW

Language permeates any given subject at any educational level or course content, and it is the means to fully understand a multitude of topics covered in school. Language can be seen as more than a mere means of communication, but also as a tool that affects and informs learners, as well as their surroundings, personal histories and futures in profound ways (Norton & Toohey, 2004). One of the most fascinating aspects of language is the concept of metalinguistic awareness (MA), or the ability to objectify language in critical and structural ways. A portion of such ability to deconstruct language is focused on grammar and the knowledge one has about grammar (Myhill et al., 2013). More important, there is an urgent need to bring the notions of MA and its grammar aspects to light through pedagogically-sound practices that will foster preservice teachers' success in their future classrooms.

MA: Why It Matters

The main aspect informing the present study is the importance of MA to preservice teachers, especially in relation to grammar. Grammar has been the subject of debate, disagreement, euphoria, and even indifference in the language teaching field. In addition, the issue of MA, or what one can objectively realize about language and grammar, is another major topic that, although admittedly relevant, has had surprisingly limited empirical investigation, especially in the teacher preparation sphere (Myhill et al., 2013).

The present study investigates MA with focus on grammar as a factor for a better professional efficiency of preservice teachers in direct relation to language, with positive

consequences for their future students, both native speakers as well as ELs. By the same token, most preservice teachers who participated in the study were native speakers of English and proficient in their language. As a result, MA in this context is not related to language acquisition or second language performance, but to the ability of deconstructing language to be better understand it structurally. Although earlier studies have brought these questions to light, the present study attempted to shed light on the specific preservice teacher aspect of MA with a grammar focus.

In order to expand the notions on MA, it is of paramount importance to look back, albeit briefly, to at least two of the most representative views from the vast body of work devoted to aspects of language acquisition and the mind, and how the issues of MA and grammar can be framed within this process. In other words, how has the study about language awareness evolved to where it is at this point, and why is it still relevant for the field of education?

Views on Innate and Behaviorist Aspects of Language

When Noam Chomsky came into the applied linguistics scenario of the 1960s, B. F. Skinner had just delivered to the world a series of lectures on his behaviorist framework on the subject of language. These lectures would form the basis for his later book *Verbal Behavior* Skinner (1957). He proposed a language framework that was largely based on human behavior, in a way almost as if language was a direct result of human behavior. Chomsky's criticism of such views was fierce (Chomsky, 1968), and he accused Skinner's behaviorist perspective of being too simplistic at explaining the incredibly creative ways by which human beings are able to both understand and produce myriad novel sentences that they had never heard before, and do so without any apparent extraordinary effort. The behavioral approach was too functional and in

17

no way attempted to explain the nature of language acquisition itself. Moreover, Chomsky argued that imitation (from children) and direct coaching (from parents) could not be considered the basis of language acquisition.

Views on Language Acquisition and Grammar

In his book *Language and Mind,* Chomsky (1968) explains that language acquisition takes place in quite independent ways, despite exposure, and that the amount of new linguistic material that human beings are capable of generating without intentional training could only be the product of the construction of an entire cognitive system, and an individual and independent one for that matter. For Chomsky (1968), a theory of grammar primarily answers a fundamental question about the *nature* of human beings' knowledge of their language, and the nature of *how* such knowledge allows them to creatively use language, mastering rules and being able to associate sounds to meaning and from there produce an infinite number of new sentences. What he calls *generative grammar* is defined by a system of rules that are constantly making sound-meaning relationships, and by doing so it produces an infinite set of abstract items of language.

Conversely, and there lies the most relevant aspect of Chomsky's elaboration on his theory for the present study, there is no *conscious awareness* of the incessant mastering of the aforementioned rules while humans acquire their first language, or of how these rules are being routinely and ordinarily used. Crucial to the discussion of the pursuit of a conscious analysis of the grammar of a language is the need or lack thereof for such effort. As a functional user of any given language and for routine communication needs, does one need to know *about* language? Is the explicit understanding of the system of rules which dictate how a language is organized something that we should struggle to achieve? In fact, Chomsky questions why this sophisticated

system of rules should be ever invited into consciousness. It is indeed a reasonable question: what purpose would it serve if the functionality of the language user has already been ensured effortlessly?

> Each of us has mastered and internally represented a system of grammar that assigns structural descriptions to these sentences; we use this knowledge, totally without awareness or even the possibility of awareness, in producing these sentences or understanding them when they are produced by others (Chomsky, 1968, p. 104).

This perspective is extremely important to the current investigation: the effortless fashion through which human beings produce and understand language. Native speakers have no overt knowledge of how their language works, and they are perfectly functional while using their languages without this awareness. It is only through an intentional desire or a need to do so that adults will consciously attempt at language analyses that are structural in nature.

Children and MA

The way MA is processed and how it can be improved in relation to children's development may shed some light on the issue of MA for preservice teachers and the need for them to be also metalinguistically aware. In addition, such understanding can prove useful by allowing teachers to better grasp what their students will face at different developmental stages in relation to MA, consequently assisting them more adequately.

Metalinguistic abilities are thought to play a crucial role in literacy acquisition in young children, as well as being of paramount importance in the transition from preschool to the years of formal schooling. However, Tunmer et al.(1984) recognize the difficulty of defining MA in more precise terms due to fact that the nature, functions, and the specific age of onset are still

largely debatable issues. This notion is espoused by Bialystok (1985), not only in relation to age of onset but also concerning which activities can be called metalinguistic. For Bialystok, the concept of MA goes beyond mental ability, encompassing a wider set of issues, all of them sharing specific traits.The MA of children has been investigated in numerous studies, both in terms of age of onset as well as concerning how to assess its multiple developmental stages in early life, not to mention the multiple types of awareness that can be separately assessed during children's development (Apel et al., 2012; Bialystok & Ryan, 1985; Bialystok & Barac, 2007; Cairns, Schlisselberg, Waltzman, & McDaniel, 2006; Lewkowicz, 1980; Moats, 2009; Zipke, 2008, 2011).

Phonemic awareness, for example, which is recognized as an essential branch of MA, has been largely considered one of the factors that will assist children in learning how to read, assisting young readers at beginning levels, and will contribute to continuous reading progress. According to Lewkowicz, (1980), the high correlation between phonemic awareness and reading success has been well established, as well as practical applications of research related to it in pre-reading training programs. A detailed list of phonemic awareness tasks includes: sound-to-word matching, word-to-word matching, recognition of rhyme, isolation of beginning, medial and final sounds, phonemic segmentation, counting phonemes, blending, specification of phoneme deletion and finally phoneme substitution.

It is clear how this understanding of phonemic awareness relates directly to the teacher in charge of implementing the aforementioned tasks: if teachers are not able to grasp their importance, they will not be able to successfully apply them. In a similar vein, preservice teachers should be exposed to the same opportunities to develop multiple aspects of their MA.

When they are not given such opportunities, they lack both the knowledge and the confidence to deal with specific language issues. Fielding-Barnsley (2010) investigated 162 preservice teachers, and found that most of them felt unprepared to teach beginning readers, especially in relation to phonemic awareness and the sound structure of words. It is interesting to note that the same teachers in this study demonstrated positive attitudes toward using phonics instruction while teaching reading, at the same time admitting that they were not knowledgeable enough to do so. Having a positive attitude toward teaching phonics is a good step, but that alone does not suffice for teachers to be able to teach to their students what is not clear to them.

Another aspect of MA, semantic awareness and the ability to detect lexical ambiguity in sentences is also meaningful in relation to children and their development. If teachers understand that MA can be fully developed if fostered through opportunities to engage with language, they can be an active part of this engagement, and expand their students' MA as well as their own. In a longitudinal study that followed children from preschool through grade three, Cairns et al. (2004), found out that participants' ability to recognize that a sentence had both lexical and structural ambiguous meanings was a robust predictor of their future reading scores in subsequent grades. The ability to understand, for example, that the sentence 'The chicken was ready to eat' can be interpreted in two completely different ways was noticed in second grade, predicting third-grade reading ability (p. 130).

Grammaticality Judgments: Relevant since childhood. Another aspect of MA and one that is directly related to the present study, grammaticality judgment is directly related to syntactic awareness, which has been shown to help young children in reading comprehension tasks. The ability to make grammatical judgments, which is also a type of metalinguistic skill,

21

has been connected to cognitive development as well as reading acquisition (Cairns et al., 2006; Bialystok & Ryan, 1985). As such an important aspect of children's development, it should also be a clear aspect of preservice teachers' expertise with their future students.

Cairns et al. (2006) investigated the metalinguistic skills of 77 children from four to six years of age, all of them native speakers of English from middle to upper classes and without any type of cognitive or linguistic impairment. Participants were given ten pairs of both ill- and well-formed sentences reflecting a number of syntactic structures, all of them appropriate for the participants' age levels. The study aimed at determining both the ability of the participants to judge grammaticality as well as to determine at what age such ability is at its most developed level. Participants were explicitly asked to *think* about language and linguistic form in order to perform the proposed tasks, a significant step towards cognitive awareness and consequently MA. The researcher engaged the children in conversations about language and modeled comparisons between the way things are said in English and Spanish, for instance. Feedback was provided for all the wrong guesses in this preliminary section, to make sure that participants were fully informed of what was expected of them during the proposed tasks, as well as to turn the preparation itself into relevant teaching moments.

This study's results suggest that the most developed group was the 6-year old one: their average scores were 15.7 out of 20, against 4.70 and 1.43 for the 5 and 4 year-old groups, respectively. These findings suggest that there is an age of onset for MA demonstrated by these language exercises, and that older children benefit more from them. Montrul (2008), on the other hand, suggested that around the age of four children start *actively developing* MA. This type of knowledge is of particular importance in the teacher preparation field in the sense that it helps

teachers better gauge what types of activities can be beneficial to their students at different stages of development.

In the psycholinguistic development sphere, children's language development can offer two distinct dimensions of awareness, namely epilinguistic and metalinguistic. In the former, the child shows intuitive awareness of the implicit grammatical rules of language, an ability that can be illustrated by the child's recognition of ungrammatical occurrences in the language. In the metalinguistic domain, not only can children recognize the ungrammaticality in a sentence, but they can also consciously demonstrate such understanding by explaining why the sentence is ungrammatical (Karmiloff-Smith, 1979). While epilinguistic behavior is demonstrated earlier in life, metalinguistic behavior will be available around the age of five years or older, echoing findings from Cairns et al. (2006)

In summary, understanding how metalinguistic abilities are fundamental in children's development of literacy is crucial to the present study in a two-fold way. First, by understanding how children start exercising their metalinguistic abilities at a very young age, and make use of them in order to learn to read and write, preservice teachers can better tap into these resources in order to help their own future students. Second, this study poses a question in relation to the apparent lack of metalinguistic abilities for adults. Why does it seem so far-fetched for preservice teachers, as in the present study, to better understand their own language in more structural ways? If metalinguistic abilities are part of a natural and beneficial stage of children's development, what happens to these abilities as children become adults? These metalinguistic abilities will be essential to preservice teachers in their future classrooms.

Metalanguage

A salient aspect of MA is metalanguage, or the ability to use language to depict, analyze or explain language (Ellis, 2005; Myhill et al., 2013; Berry, 1997; Wray, 1993). Metalanguage and its implications to preservice teachers are also at the core of the present study. It is very common for language users in general to feel unable to clearly explain grammatical issues with proper language. Oftentimes they are able to correctly detect ungrammatical formations but are unsuccessful at clearly stating what is wrong with them, and even less successful at using proper grammatical terms to elaborate on the issue at hand (Clapham, 2001; Alderson et al., 1997).

This same inability is surprisingly also very common for teachers, even the seasoned ones (Borg, 2003; Moats, 1994). Although they are able to point to errors with confidence, that confidence quickly fades when they cannot find the right word to name grammatical terms. While that is acceptable for the regular language user, it should not be the case for teachers. According to Folse (2016), when native speakers are confronted with a specific language inquiry or issue, they rarely know why their L1 functions that way and they tend to rely on three vague attempts at explanations. These attempts are: a) the issue at hand might be an exception to the rule; b) it is right because if it sounds right, it probably is; or finally c) it is right because that is how native speakers say it. Understandably, such empty and somewhat obtuse explanations are in no way an acceptable approach to helping a student with a grammar question, whether that student is an English learner or native speaker of English. None of these explanation attempts is good teaching. Teachers without expertise on essential notions about English as a second language should not be in the classroom because they do not have the requisite knowledge to help ESL learners (Folse, 2016).

Another important point to be considered is that teachers can choose to use metalanguage minimally or not at all, depending on their expertise to gauge how necessary that use is in specific educational circumstances (Andrews, 1999b). In some cases, a decision to avoid metalanguage use in favor of another approach to dealing with a grammar issue can be a sound solution (Alderson & Hudson, 2012; Berry, 1997; Harper & Rennie, 2008). However, there is a difference between not using it as a pedagogically-sound and conscious decision and not using it at all for lack of *knowledge* about it. It is crucial that teachers know the basic terminology to talk about language in technical terms (Alderson & Hudson, 2012), extrapolating vague elaborations such as "all these little like name things" when referring to nouns, verbs or prepositions (Andrews, 1999b, p. 147). In other words, knowing the proper terms to define and explain grammatical issues should not equate to drilling students with an infinite list of difficult, technical terms and overwhelming grammar explanations. It is essential that teachers are metalinguistically aware but also careful in relation to when, how and how much metalanguage to use, and how to make it as pedagogically-sound as possible.

As stated before, lack of proper metalanguage while analyzing language has been demonstrated in studies' findings related to teachers (Berry, 1997; Williamson and Hardman, 1995), which might pose a question: if the teachers themselves are not able to correctly employ terminology about language, how can students be expected to develop it?

MA is Teachable

Numerous studies have clearly demonstrated with their result that MA can be taught, learned and enhanced through conscious efforts and adequate training. In the context of the present study, this claim is being made both to first as well as second language, in the sense that

MA is relevant and beneficial for both aspects. The investigations previously mentioned in Cairns et al., (2004, 2006) are particularly important to the scope of the present study because their findings show that metalinguistic skills are not only a natural development in children's language development, but that they are also *teachable*, and that positive results can be achieved in a rather small amount of time.

In discussing the study's results, Cairns et al. (2006) emphatically addressed the fact that the instructions provided in the beginning of the study were the main facilitator of the performances because it created "a mindset to regard language as an object of consideration and analysis to the extent that they were cognitively able to do so" (p. 217). This perspective is espoused by the current investigation concerning the fact that not only language *learners* may be invited to better understand the intricate aspects of language, but also already proficient language *users* of any age should be invited to do the same, especially if doing so can bring specific benefits to their professional goals.

In a study with 46 third-graders from low socio-economics and multiple cultural backgrounds, Zipke (2008) found that with only two hours of instruction in the identification of homonyms, sentence ambiguity and riddles, participants' MA was increased. In addition, the students who took part in the study were positively engaged and enthusiastic about learning, performing significantly better from pre- to posttest. The same author investigated the issue of homonym detection again in a study with 36 first-grade participants, also with a pre-and posttest (Zipke, 2011). Participants received MA instruction and demonstrated not only improvement from pre- to posttests, but were also deemed more flexible in their ways of thinking as well as more open to alternative solutions to interpret the proposed texts. The study's findings pointed to

two fundamental aspects of MA and instruction: (a) MA can be a trainable skill, and (b) students can be both enthusiastic and successful while being exposed to activities that aim at increasing their metalinguistic skills (Zipke, 2011).

In another study with 56 second and third-grade children, Apel et al. (2012) investigated the phonemic awareness, morphological awareness, orthographic awareness, receptive vocabulary, and rapid-naming abilities of participants, and determined how these abilities predicted the children's reading and spelling skills. Regression analyses revealed that morphological awareness was the sole contributor to spelling and, together with orthographic awareness, uniquely contributed to word recognition. Morphological awareness was also significantly related to reading comprehension. These results add to a growing literature base providing evidence that early literacy development is influenced by morphological awareness, an ability that has received considerably less educational attention, especially if compared to its phonemic counterpart. Additionally, the findings point to the importance of tapping into multiple sources of metalinguistic knowledge when providing instruction in reading and spelling (Cairns et al., 2004 2006; Moats, 1994, 2009; Zipke, 2008, 2011).

The Role of Explicit Knowledge for Teachers. While inviting teachers to better understand language, the idea of dealing with language in either implicit or explicit ways comes to mind. In his book *The Architecture of Language*, Anderson (1983) addresses the major issue of defining if language acquisition happens in the same lines of other cognitive processes. According to him, all the higher-level cognitive functions, among them language, occur under the same basic pattern, or what he calls "the same underlying architecture" (p.261). Although not designed in relation to language exclusively, one of his main contributions to the notion of how

language is processed refers to a perspective that is similar to the implicit and explicit notion of language production.

According to Andrews (1997, 1999a, 1999b), in order to be an effective and successful teacher, one needs both the implicit and explicit notions of language knowledge, and the capacity to draw on both, depending on the circumstances. He also points to the differences between notions of procedural and declarative knowledge. Although Anderson's (1983) views were initially related to cognitive psychology, they have been used as vital points while describing the ways by which a teacher (a) knows about language, and (b) can demonstrate this knowledge and apply it to concrete situations, either using metalanguage or not (Andrews (1997). Ellis (2006) points to the fact that implicit knowledge is unconscious, procedural and not verbalized, while explicit knowledge is held consciously, being both learnable and verbalizable. Gregg (1989) also makes a clear distinction between the notion of *knowing that* and *knowing how,* with the former implying a kind of knowledge that happens by chance in a natural, spontaneous way, or what Ellis (2008) calls primary knowledge.

All these notions are important while addressing the need for preservice teachers to expand their language abilities to be able to handle language inquiries, from their future students. Preservice teachers must know more about language and be able to gauge when and how to use this knowledge depending on specific teaching scenarios. However, and more important to the discussion concerning grammar and teaching, is the fact that the language knowledge of the preservice teachers in the present study is not being framed in the regards to proficiency, as Krashen posits (1981). Most preservice teachers depicted in the present study are native

speakers, proficient users of the English language. Consequently, the debate between explicit knowledge not becoming implicit for the purpose of an enhanced proficiency is irrelevant.

What matters for the proposed discussion of language knowledge and MA for preservice teachers is quite the opposite: What can be done once implicit knowledge is already established, but a lack of structural understanding is demonstrated? Can *analyzed*, as in *explicit knowledge* have a positive role in how teachers understand their own native languages in more structural ways? More important, can this knowledge positively affect how preservice teachers will relate to their students' language struggles?

Notwithstanding, no matter the theoretical inclination one might have regarding implicit and explicit types of knowledge, the notion being espoused by the current study is one that takes into consideration the reason why explicit knowledge should be a part of preservice teachers' preparation: teachers who are better prepared language-wise will be able to offer better language guidance to their future students.

Teacher Preparation and Grammar

The present study is not differentiating the need for grammar as more necessary for L1 or L2, but advocating it for all preservice teachers, independent of what and who they are going to teach in their future careers. Pre- and inservice teachers should have a more robust grasp of grammar and MA, regardless of the population they are going to teach (Andrews, 2003; Berry, 1997; Borg, 2003; Celce-Murcia, 1992; Moats, 1994, 2009). Language knowledge and grammar expertise (or lack thereof) will affect preservice teachers' future careers, as well as their preparation to become teachers. The teaching of grammar for preservice teachers is not only relevant but essential for their future careers (Kolln & Hancock, 20005).

One of the most recurring questions vis-à-vis grammar and instruction is that of its importance or lack thereof. Grammar and grammar teaching, for that matter, have become a polarizing topic in the field of applied linguistics for many decades now. There was a time when teaching grammar seemed to be an undeniable necessity at the core of the language teaching and learning dyad. Not even audio-lingual methods of language instruction, which had become very popular in the 1970s, were able to eliminate grammar instruction for second /foreign language teaching. Students were required to memorize prescriptive rules and know extensive grammatical terminology (Bloor, 1986), which seems as distant from current pedagogical trends as the outdated translation method.

A lot has changed since then, and it should have, for grammar should not be taught in a decontextualized fashion, disconnected from the specific pedagogical needs of a given group of learners. Celce-Murcia (1992) emphatically addresses the same notion in an article discussing the merits or lack thereof of teaching grammar:

> …I must point out that grammar instruction carried out
>
> for its own sake, totally divorced from activities that involve using it as a
>
> resource to convey meaning is as irresponsible and counterproductive as
>
> not teaching grammar at all. (p. 408).

This is a point of view thoroughly shared by the current study's perspective. In addition, it is important to note that the present study does not advocate explicit grammar instruction for all and without a solid pedagogical understanding of its need or lack thereof. Moreover, it does not endorse the old model of drill-oriented practices to preservice teachers; neither does it suggest

the same for their future students. As stated by Ellis (2006), obsolete methods of grammar that have failed numerous attempts to show effectiveness in the classroom are not a viable option.

Should Grammar Be Taught? When? Although the importance of grammar in relation to preservice teacher preparation is considered undeniable within the scope of the present study, the fashion through which grammar infusion should and could be done, if deemed necessary, should follow clear pedagogical needs that vary from grade to grade, as well as take into consideration the many different educational settings and individual variations at play.

Celce-Murcia (1992), states that there are "definitely circumstances where formal grammar instruction is necessary" (p. 408). The challenge then is to be able to gauge when and for which specific educational scenarios that formal instruction applies. Moreover, it is essential that language knowledgeable teachers are also equipped with the pedagogical understanding of how to approach grammar instruction in meaningful and effective ways to their students, since the idea of grammar teaching should not imply one format of instruction that will address and benefit learners in all possible scenarios.

Classrooms with language learners, learners with disabilities or simply learners that are not at the same level of progress as their peers should receive individual consideration in relation to what can be done in terms of language support. By the same token, classrooms where pre-literate learners primarily aim at basic oral skills, or where children are being exposed to their second language under optimal learning environment, a more formal and intensive grammar instruction is perceived as neither necessary nor beneficial (Celce-Murcia, 1992).

Negative Views on Grammar Teaching. For some, the fact that what is called *traditional grammar* for many languages is primarily based on Latin and Greek, and that such

basis was engineered with minimal modifications, should be reason enough to discredit grammar teaching as outdated and impractical (Henshaw, 1985).

By the same token, if the idea of communicative goals is of essence, why should methods that emphasize grammar structure be considered valid in the current educational scenarios? Should grammar be taught, and if the answer is yes, how should it be taught? Not only that, how much and how often should grammar be taught? In fact, some question if grammar should be taught at all, depending on some circumstances (Ellis, 2005; Henshaw, 1985; Krashen, 1992, 1993). For some linguists, grammar instruction has never been relevant enough to be explicitly taught. A devoted opponent of grammar instruction in the SLA sphere, Krashen (1981, 1982, 1992, 1993), has consistently pointed that explicit instruction of grammar does not lead to deeper levels of knowledge, or what he called an *acquired competence* for second language learners.

In the same vein of Chomsky's non-conscious, effortless approach to language acquisition for first language, Krashen developed the influential concept of the Monitor Theory in relation to second language acquisition. According to this theory, a conscious knowledge of grammar can only be available and accessible as a type of editor, which can only function under three distinct circumstances: (a) learners need to know the proposed rule, (b) learners must have time enough to apply the rule, and finally (c) learners need to be intentionally alert to the form of the language structure at hand (Krashen, 1982). According to Krashen, only under such circumstances can higher accuracy be obtained but at the cost of a speech that is not as fluent and fast as that without using the editor.

Krashen (1982) has always defended the position that the ability to use a language without conscious effort on the part of the learner is not dependent on any explicit learning; in

fact, the excessive use of the so called Monitor can be anything but beneficial, as it inhibits learners' performance, somehow crippling what should be a spontaneous process. In an article commenting on the findings of a study on grammar instruction, although the overall results seemed to suggest that there was effective learning of the grammatical topic, Krashen (1993) dismisses the positive outcome as different from *real acquisition*, but simply a result of the over drilling of one specific item to the point of exhaustion for the students: "The effects of grammar teaching still appear to be peripheral and fragile" (1993, p.725).

Positive Views on Grammar Teaching. However, the rejection of grammar as a reason to privilege communicative goals is neither necessary nor beneficial to both students and teachers. As stated by the National Council on Languages in Education (NCLE):

> It is also time to point out that the effort of getting rid of the old style of grammar teaching required the exclusion of most investigations of pattern in language, and that a lot was therefore lost (NCLE, 1985).

In reality, the criticism toward grammar instruction might be more a matter of how it is being framed than a fair assessment of the need for it. It seems that more often than not, discussions about the merits of grammar teaching tend to revolve around definitive answers as opposed to a more inclusive perspective that is open to all lines of inquiry. Such an extremist approach is counterintuitive to any educational agenda: grammar should be where there is a need for it, and that should not eliminate or decry any communicative trend.

Ellis (2006) points to the fact that grammar teaching has been commonly associated with a mere presentation and practice of the many grammatical items available in any grammar book. Although this format can be an option depending on circumstances, it is not the sole possibility

in terms of grammar instruction. Grammar teaching can be done with mere exposure to multiple input samples as they illustrate language structures. In other words, and central to the discussion proposed in the present study is the notion that grammar does not need to imply a monolithic, rigid model, but it can be adapted to different needs and within multiple pedagogical objectives. Moreover, there is a difference between discussing the need for grammar instruction on the part of learners and the same need on the part of *any* educator. Teachers in charge of learners' language instruction decisions should be better equipped with grammar knowledge, once again regardless of their students' status as either native or nonnative language speakers.

Equally important, the main issue depicted in the present study does not concern the concepts of *acquiring* versus *learning* of a second or even first language, and how grammar can affect, benefit or halt progress for each of these processes. There is no ideal answer to questions on why and how grammar should be taught or not, in as much as the vital implication about these questions is that there is no single answer to them. Adaptations should be made taking into consideration specific pedagogical needs, and concerning preservice teachers' needs there is a strong research body that points in a positive answer to grammar teaching.

The main consideration at the core of this discussion relies on the fact that when it comes to educational settings, language is probably the only common element that permeates all the other subjects, at any level of difficulty, and for all ages, hence the need for knowledge about language on the part of the teachers to be way beyond peripheral, but deep and meaningful.

Inviting a More Balanced Perspective on Grammar. No matter the fact that some insist on the unfair and exclusive all-or-nothing approach to grammar, there is a wiser point in-between the two extremes, where dismissing rigid explicit grammar instruction *does not* equate

to dismissing grammar knowledge and expertise completely. Crystal (2006) insists that although grammar does not encapsulate all there is to know about language, it is of critical importance in dealing with language. While depicting the more modern and meaningful approaches to the British A- level English classes nowadays, he describes:

> You would see students looking critically at the words people use, the sentences in which they use them, the way in which the sentences are put to work in discourse and whether these discourses suit the context in which the speakers or writers are operating. In short, they are learning to judge appropriateness in others.
>
> …We can sum it up in another way: students are being taught to recognise and understand the consequences of making linguistic choices (Crystal, 2006, p. 210–211).

It would be challenging to describe the combination of language analysis and fun in better words. In the same vein of Zipke's (2008) description of how children can find language fun and entertaining, so can adults, if it is presented to them in robust but nonetheless, interesting ways.

Teacher Preparation and MA

The point at which MA and teacher preparation come together and why this fusion is being proposed in the current study is simple: there is a crucial difference between the need for enhanced language knowledge for a professional in non-educational areas and that of preservice and inservice teachers in general.

Studies' findings consistently point to the importance of myriad interventions for students that will rely heavily on their *teachers*' expertise in numerous content domains, such as phonology, phoneme/grapheme correspondence, morphology, semantic organization, syntax, discourse, and pragmatics (Cairns et al., 2004, 2006; Moats, 2009, 1994; Zipke, 2008, 2011). Nonetheless, if the professionals who are expected to conduct such interventions are not knowledgeable and confident enough, it is fair to expect that these interventions will not be carried out successfully.

However, for decades now the primary aspects of MA, in relation specifically to grammar teaching have been addressed in studies that depict SLA scenarios, whereas little has been done in relation to grammar teaching for first language acquisition and teachers' knowledge of language.

Borg (2003), states:

> For over 20 years the predominant source of knowledge about grammar teaching had been studies of second language acquisition (SLA), where the focus was on learners and learning outcomes; in the work reviewed here, we see evidence of a conceptual shift which is based on the realisation that SLA research has not provided definitive answers about grammar teaching.

In a seminal study that investigated the merit of grammar teaching, Bloor (1986) administered a questionnaire named SPAM (Students' Prior Awareness of Metalinguistics) to 63 students who were about to start Modern Languages and Linguistics degree courses, and to 175 second-year students of other degree courses, making a differentiation between the two groups as linguists (the former), and non-linguists (the latter). This was arguably the first of many instances

36

in which this questionnaire was used and adapted by researchers to test language knowledge and MA. The questionnaire contained traditional grammar items such as parts of speech and parts of sentence (named grammatical functions in the original test), in a recognition-based format: they had merely to recognize the parts of speech from sentences. Besides these questions, participants were also asked to give an example of one way in which "English differs grammatically from some other language" (Bloor, 1986, p. 160).

The results from this study indicated a reason for concern with the state of ignorance of 'school-leavers' and college students with respect to grammar and language awareness. Although a thorough discussion about all the questionnaire items as well as its results has no merit at this stage of the current study, it is relevant to say that out of all the parts of speech available, the only two correctly recognized by the linguists were *verb* and *noun*. In the non-linguist category even these two parts of speech failed to be named correctly. Bloor (1986), citing the surprising overall ignorance that the results showed about students' MA, argues that although it is true that communication goals at the time seemed to have taken over the language instruction realm, there should not be a need to *choose* between communication and grammar, but to embrace both, in different ways. Crystal (2006), points to the fact that teacher cognition and the instructional choices that teachers make are fundamental in order to address the gaps in teachers' understanding of grammar.

Could this lack of research on teacher cognition suggest that institutions believe that preservice teachers, especially the ones who are native speakers of English, are perceived as good functional users of the English language, and as such do not need any extra exposure to explicit language instruction? Harper and Rennie (2008), point to a clear gap in metalinguistic

knowledge at the time of graduation for preservice teachers in the United States, while comparing this reality with a different, more grammar-aware approach in the United Kingdom.

Conversely, the core issues that permeate MA for teachers are essential to their future students and their language development. Teachers who are more linguistically aware will be able to better gauge and understand both the language they teach and their students' language difficulties (Fielding-Barnsley & Purdie, 2005, Fielding-Barnsley 2010; Wright, 2002).

Mainstream Teacher Preparation. The needs and remediation options for students struggling with language issues, whether in relation to literacy or to ELs' more specific language needs, have been reason for concern and research in many lines of investigation. However, at the core of this issue are the professionals who are supposedly well-versed in all aspects of language, and who can provide assistance to students: their teachers. But can they do so?

In relation to mainstream teacher preparation, Moats (2009), argues that in order to understand the needs of a diverse population of students, teachers themselves must have a robust knowledge of language structure, as well as reading development and pedagogy. Studies have shown that teachers must be knowledgeable in multiple aspects of language, among them phonology, morphology, semantics, syntax, discourse, and pragmatics in order to assist their students in their development (Fielding-Barnsley & Purdie, 2005; Fielding-Barnsley, 2010; Moats, 1994, 2009; Myhill et al., 2013). However, little empirical research has been done addressing the role of grammar teaching and grammatical understanding for L1 learners (Myhill et al., 2013).

In a recent study that also discusses the need for more empirical research on language knowledge and MA, specifically toward grammar and its impact on writing, Myhill et al.,

(2013), drew on a subset of data from a previous larger study to investigate 32 teachers with their respective students totaling around 744 students at ages from 12-14 years old. The study investigated if grammar teaching embedded within regular writing teaching units would improve students' writing attainment. Results pointed to a strong impact of the intervention, but even more important, the results also demonstrated that the *teachers*' subject knowledge of grammar was a mediating factor in the success of the intervention. This study included 96 class observations, individual teacher interviews, and one focus student interview from each class among the data collection procedures.

The study above was outstanding in terms of the number of participants and the scope of investigations, drawing from both quantitative and qualitative data. The authors concluded their discussion by addressing the need to develop teachers' grammatical pedagogical content knowledge. Grammar knowledge without the knowledge of how to use it can be deemed insufficient to assist students, a notion that Andrews (1997, 1999a, 1999b, 2003) has explored extensively in his grammar teaching discussions, studies and articles, under the terminology procedural versus declarative knowledge. Declarative knowledge goes beyond mere knowledge about grammar; it refers to knowing what to do with the knowledge itself.

For example, concerning the ability to discern between what one knows and how one uses that knowledge, Andrews (1999b) emphasizes the need to know when *not* to make use of explicit knowledge. In other words, having MA does not mean drilling students with complicated explanations or terminology. Teachers can have, for example, a very good knowledge of metalanguage, but they should weigh their options of either using their technical language knowledge in the classroom or not, depending on the individual, specific circumstances that are

presented to them, pedagogically speaking. Using an intricate, highly sophisticated list of grammatical terminology is not necessarily good practice under all circumstances. In reality, it can be both inefficient and discouraging to students.

According to Moats (1994), graduate level teachers are clearly ill-prepared to deal with the really demanding task of teaching reading and spelling explicitly. In a study that tested inservice teachers and their language awareness, including current knowledge of terminology about morphology, phonetics, phonology and phonics, the results were "surprisingly poor", showing that even when teachers are young and motivated to perform to the best of their abilities they lack the basic foundations of the language structures in order to assist their students in those areas (p.81). Where then, should they have acquired those basic foundations if not during their preparation years as preservice teachers?

According to Fielding-Barnsley (2010), an ideal balance of oral language, phonemic awareness, phonics, word identification, fluency and vocabulary, among other factors, can provide teachers, both preservice and inservice, with the necessary knowledge and support they need to be of better assistance to their pupils. It has been well-established in reading research that phonemic awareness is of great value for beginner readers, and it also predicts successful reading development in further grades. Nevertheless, teachers must possess explicit knowledge of these metalinguistic skills themselves in order to help their students (Moats, 1994, 2009, Fielding-Barnsley & Purdie, 2005; Fielding-Barnsley, 2010).

In a rare article in which three aspects of grammar instruction were discussed, Borg (2003), looked at studies on first, second, and foreign language classrooms and the issue of teacher cognition and grammar declarative knowledge, and what was known in the United

40

Kingdom as Knowledge about Language (KaL). In the same vein of Andrews (1999a), the idea of deficient grammar and metalanguage knowledge on the part of teachers was considered alarming. If teachers' knowledge of language is so precarious, how can they be prepared to help their future students, who are arguably at lower levels of language knowledge?

Wray (1993) investigated student teachers and their ability to identify, for instance, adverbs, pronouns and prepositions, arguably basic parts of speech that any educator should be able to easily decode. However, participants achieved mean scores of respectively, 30% (for adverbs), 23% (for pronouns) and less than 10% for preposition identification. Such findings resemble Bloor (1986) and his investigation on the same aspects of grammar, with similar low results. Williamson and Hardman (1995) looked at 99 trainee primary school teachers, all of them beginning a postgraduate certificate in education. On a question asking them to name *parts of speech*- arguably a very basic topic in the language knowledge hierarchy, participants achieved a mean score of 5.6 out of 10, with performances at their weakest on items asking them to label different kinds of clauses. As they analyzed the study's findings, researchers pointed to the abysmal gap in participants' knowledge about grammar, as well as their numerous misconceptions about language.

Moreover, the study's findings suggested a clear lack of proper metalanguage for analyzing language use, an issue that was also investigated in a study with 372 undergraduates in Hong Kong (Berry, 1997). In this study, an interesting aspect of language awareness was proposed: the awareness of the teachers with reference to their students' knowledge. In other words, the author poses a question on the importance of teachers to be able to gauge how much their students know about metalinguistic terminology, and to adapt their practices to match this

knowledge. The study's results showed clear discrepancies between what the students knew and what their teachers expected and had predicted that they knew.

Anecdotally, the same lack of knowledge about grammar terminology and more specifically about parts of speech has been noticed in the applied linguistics course attended by preservice teachers in the present study. Parts of speech are essential in any primary analysis of a simple sentence, yet most preservice teachers in the class are openly at a loss when asked to name parts of speech or to identify them in sentences. This issue that was clear in a pilot administration of a similar version of the original Language Awareness Test conducted in the fall of 2014 by the principal researcher of this study. When preservice teachers nearing graduation and even inservice teachers find it challenging to identify a noun or understand the difference between a noun and a verb in a sentence, there is reason for concern, as well as a need to act upon it.

Although research on teacher MA seems to be mostly geared toward L2 teachers, these research findings are valid for any teacher preparation analysis concerning language. A long-standing advocate for an enhanced grammar curriculum that addresses teachers' language deficits, Andrews (1997, 1999a, 1999b) has made constant use of the term Teacher Metalinguistic Awareness (TMA) for teachers' knowledge about language. Admitting that grammar had suffered a decline in positive views about its importance in earlier decades, especially due to the communicative language teaching (CLT), he stated that by the end of the 90s grammar started to return into the teaching scene, a point of view shared by Alderson and Hudson (2012).

The need to draw on what teachers know and how they can use that knowledge was also espoused by Galguera (2011) in a proposal to define teacher preparation in a wider scope. Instead of teacher preparation for a specific 'type' of student, namely ELs, which he considers "too broad and not inclusive enough" (p.86), Galguera points to a conceptual framework that considers the complexities of language, especially for academic development. Preservice pedagogy must be questioned and brought to light in order to better understand the teaching-learning relationship, building upon the concept of pedagogical content knowledge, or a common ground between content knowledge and pedagogy.

Studies suggest that preservice teachers feel much trepidation concerning their preparedness to deal with language learners. In a longitudinal study that investigated the beliefs of 130 content-area preservice teachers, Pappamihiel, (2007) collected journals containing reflections for over three years. Every semester, a total of 10 weekly journals per participant were collected as data, and were analyzed throughout the semester; participants also participated in a type of community-based service learning with ELLs. Among the emerging topics in the journal there were fears and concerns related to working with ELLs, as well as a discussion on the legitimacy of doing so. Preservice teachers in this study also expressed a variety of negative feelings toward teaching ELLs, among them indifference, misconceptions about immigrants and even a perception that their future ELL students will represent "an extra burden on their time" (Pappamihiel, 2007, p.55).

According to many of these journal reflections, content-area preservice teachers had a clearly distinct perspective on the knowledge they have of their content area and the need to understand language development on the part of their ELLs. However, at the end of the project,

the results showed a positive change in the attitudes and beliefs of preservice teachers. Although the author voices the limitations of the study, among them the pressure to report positive change from the part of the preservice teachers, the results point to a need to address these issues and the belief that, although they are admittedly difficult to change, progress can be made if they are at least seriously taken into consideration.

This study also pointed to an important issue related to language, language development and preservice teachers: the intertwined aspects of content area topics and overwhelming need for language proficiency to progress in any of them. No matter the topic, language will permeate any teaching, and MA will promote or hinder success in any learning attempt.

Context of the Present Study. The investigation proposed in this study looked at preservice teachers (mostly Education majors) from a teacher education program at a large university in Florida. Among their mandatory courses, these preservice teachers are offered two specific state-mandated courses that primarily aim at raising their awareness about second language acquisition issues, largely due to the great number of K-12 students who are not native speakers of English in the state of Florida. Preservice teachers are expected to meet TESOL standards set by the state in order to obtain their ESOL Endorsement.

One of the five domains in these standards is Domain 2: Language and Literacy (Applied Linguistics), which focuses on the L2 knowledge teacher candidates must demonstrate in order to obtain the aforementioned ESOL Endorsement. Specifically, one substrata of this domain states that "Teachers will demonstrate understanding of language as a system, including phonology, morphology, syntax, semantics and pragmatics; support ELs' acquisition of English in order to learn and to read, write, and communicate orally in English." (Standard 1: Language

as a System). In other words, teacher candidates are expected to have extensive language knowledge, and must develop MA, particularly in the areas of phonology, morphology, syntax, semantics, pragmatics, and grammar in order to meet the criteria set by a standards-based curriculum for their eventual teaching certificate.

Because the aforementioned course has multiple requirements, the time spent on language knowledge and MA is very limited, which poses a limitation related to how well teacher candidates are being prepared for their future diverse student population, as well as for native speakers of English who are struggling with language issues.

Conclusion

Teacher preparation does not end when preservice teachers leave college and start their career journeys. Genuine, committed teachers are constantly looking for ways of improving, altering, and adding to their practices, and doing so is as much part of their routine as what they already do with expertise in their classrooms. However, it is also true that the teacher preparation years during college can offer a solid knowledge base to these future professionals, which will help them in the transition from *preservice* to *inservice* teachers with success.

It was in this light that the present study expects to offer an initial picture of one of the most relevant aspects of teacher preparation: language knowledge. Although the whole idea of language knowledge is too wide a concept to be fully investigated in one single study, it was possible to assess some aspects of what was being done and how effective it was in specific relation to preservice teachers. By looking at the current state of MA in relation to grammar of a group of preservice teachers, as well as investigating if the amount of instruction they receives in

a language module was enough to alter this initial state in statistically significant ways, this

investigation looked at preservice teachers and the preparation they received in MA.

CHAPTER 3:

METHODOLOGY

The methodology for this study was based on the six interrelated principles of inquiry of any given scientific investigation, including scientific inquiry in education, as proposed by the National Research Council (Shavelson & Towne, 2002). These principles are: (a) the proposed questions should be significant and empirically investigated, allowing for testing and refuting of results; (b) the inquiry should establish a relationship between research and theoretical constructs; (c) the utilized methods must allow for objective investigation of the proposed questions; (d) the reasoning supporting the inquiry should be objective and coherent; (e) the results must allow for both replication and generalizability to other studies in the field; and (f) the research must be open to scrutiny and critique.

Research Questions

1. Is there a statistically significant increase in the overall metalinguistic awareness of participants after a treatment, as measured by pre- and post-administrations of the Adapted Language Awareness Test (ALAT)?

2. Is there a statistically significant difference in the ability of participants to identify *parts of speech* after a treatment as measured by pre- and post-administrations of the first part of the ALAT?

3. Is there a statistically significant difference in the ability of participants to identify *parts of sentence* after a treatment as measured by pre- and post-administrations of the second part of the ALAT?

4. Is there a statistically significant difference in the ability of participants to identify and correct *grammatical errors* in a sentence after a treatment as measured by pre- and post-administrations of the third part of the ALAT?

5. Is there a statistically significant difference in the ability of participants to provide *metalinguistic explanations for grammatical errors* after a treatment as measured by pre- and post-administrations of the fourth part of the ALAT?

Research Design

The research design for this study was a quasi-experimental design, *One-Group Pretest-Posttest Design* (Shadish, Cook & Campbell, 2002). The study investigated the current MA and grammar knowledge of participants as measured by an instrument called ALAT, during a face-to-face, undergraduate applied linguistics course in which they were enrolled at an urban research university in the United States.

The independent variable was the treatment, a three-week grammar module in which participants took part during one of their courses. The dependent variable, which determined the effect of the treatment or lack thereof, was the MA of participants in this study, or more specifically the scores obtained in two tests (pre- and post-) that measured their metalinguistic awareness, before and after the treatment (Lomax, 2007).

The study started on the second month of spring semester classes and consisted of a pretest immediately followed by a treatment in the form of a three-week grammar module. This treatment was a regular portion of the course in which participants were enrolled. Immediately after the last day of the treatment, participants were administered the same initial pretest as a posttest.

The aforementioned applied linguistics undergraduate course is one of two mandatory courses that preservice teachers must take to be considered ESOL- endorsed. This endorsement is part of state regulation for college majors aimed at preparing teachers in the state of Florida, and it lasts 16 weeks every spring and fall terms; it focuses primarily on issues in second language acquisition. According to its syllabus, this course is described as an examination of English phonology, morphology, syntax and semantics to assist future teachers to understand and develop ESOL instructional strategies, by equipping them with pertinent knowledge about both linguistics and first and second language acquisition research.

This state-mandated course is designed for a specific population of future teachers and its main purpose is to provide preservice teachers with overall knowledge about linguistics, research findings on first and second language acquisition, and exposure to language issues that permeate the ESOL sphere. Altogether, this course is comprised of 15 topics: language and the brain, first language acquisition, second language acquisition, phonetics, phonology, etymology, morphology, grammar, syntax, semantics, pragmatics, discourse analysis, regional variation, social variation, and culture. Since the course contains a great number of diverse topics, roughly less than three weeks are used to cover grammar and its subareas syntax and morphology. Only this part of the whole course will be examined for its impact on grammar knowledge and MA. In a more specific aspect related to language knowledge, one of the course objectives is to ensure students will be able to identify parts of speech with 100% accuracy.

Participants

The process of sampling for this study was a non-probability sampling, and more specifically a combination of convenience and purposive sampling (Dörnyei, 2007). Participants

in this study were chosen based on several factors that reflect the population of interest: (a) all of them were undergraduate students; (b) all of them were registered for an applied linguistics course which had as its primary goal raising their awareness about language issues in general, but more specifically geared to issues in second language acquisition; and finally (c) they were conveniently grouped together and available for this investigation. All participants were drawn from an urban research university in the United States, and they were enrolled in an applied linguistics course as part of their teacher preparation curriculum. Their participation was not mandatory.

Main Group and Subgroups. Three main identifiable groups of students attended the aforementioned applied linguistics course, and therefore were included in this sample to improve the heterogeneity of the sample. It was expected that most of the sample will consist of undergraduate students majoring in education, and for the purpose of this study this subgroup was called *preservice teachers*. The second group consisted of undergraduate students from TEFL certificate courses. Although these students are also pursuing a teaching career, they are not being prepared within a formal College of Education curriculum, but were in many cases pursuing teaching experiences overseas. This second group of undergraduate students take the applied linguistics course as one of four courses that will certify them to teach English as a foreign language overseas. A third and certainly minimally represented group attending this course was that of students taking this course as an elective course, without necessarily aiming at teaching careers of any kind.

A brief biodata questionnaire in the pretest addressed individual characteristics of participants such as age, gender, major, first and second language. This kind of biodata is also

relevant in terms of construct validation, ensuring that the characteristics of the group taking the test are clear and that it can be properly compared to other groups taking the same test in future studies. All participants were 18 years or older (100%); more than 84% of participants were female students; all the participants (100%) declared that they spoke English, and 18% stated that Spanish was their second language; languages named under an 'other languages' slot were Arabic, Burmese, Creole, Farsi, French, Japanese, Malaya Malan, Portuguese and Vietnamese. More than 80% of the participants were majoring in Education, and under that scope they were elementary education (52.8%), early childhood (17.3%), and finally English language arts (11.5%) major. Other majors included advertising, history, interdisciplinary studies, international and global studies, marketing, political science and psychology, totaling around 18% of the participants. Some of the participants were taking or had taken a specific course that directly addresses grammar as part of their TEFL certification. According to the biodata questionnaire, 24% of the participants belonged to that category.

Four sections of this applied linguistics course are offered in a face-to-face mode. Each section of the face-to-face course has a maximum capability to sit 32 students. As a result, a maximum number of students in the four sections could have reached up to 128, which represented the whole population of students of the face-to-face applied linguistic course in which participants are enrolled.

According to Krejcie and Morgan (1970) the sample size expected to provide an acceptable number of participants for this study would have been 92 participants, taking into consideration the sampling frame number of 128 students available. However, a decision to use the entire sampling frame of as many students as possible was made in order to ensure that

possible attrition would not jeopardize the final number of participants, consequently affecting the power of the study. Since it was determined that the whole population of the four face-to-face courses taking place on campus would be used, the recruitment was done throughout the classes, since all the students registered for the applied linguistics course were given the opportunity to participate in the research study. Therefore, the initial sample size (N=128) for this study represented the entire sample of students registered for the applied linguistics course in face-to-face mode for spring 2015. Only participants who took both pre- and posttest were considered for the final data. After accounting for all participants who were complete cases (pre- and posttest) ,the final number of participants was 101.

Internal Validity

According to Shadish et al. (2002), although the *One-Group Pretest-Posttest* design is suitable for this this study, it may still be affected by certain threats to internal validity, such as: History, Testing, and Regression.

The first uncontrolled variable is that of *History*, and it encompasses the notion that any time passed between the pretest and the posttest may produce events that affect the final outcome of the study. The longer the time passed between the two tests the higher the threat to the study's internal validity. In order to control for that, the time passed from the pre- to the posttest was as limited as possible. The pretest was administered immediately before the beginning of the morphology chapter of the grammar module, and the posttest was administered immediately after the end of the grammar chapter of the module. *History* also relates to the ways in which the two tests and treatment were conducted. For instance, any change in the way the two events (tests) were administered could become a reason for any attested difference. If one of the

instructors teaching the course decided to offer an extra grammar activity to one group and by doing so he or she managed to instill enthusiasm in his or her students in relation to the topic, that particular group could arguably present better results than the other three. In order to control for this kind of interference, instructors were required to restrict their lessons to the available materials and chapters of the book and little or nothing else.

Testing is a variable that might affect the study's results. It is predictable that, since the pre- and posttest are one and the same, when students take the posttest for the second time they will perform better. Because participants are being exposed to the same instrument for the second time, they have a better adjustment to the idea of taking the test again, as well as what is expected from them. Another factor that can be taken into consideration is the fact that, since participants are classmates taking the same course, they can communicate with each other after the first test and exchange ideas and opinions about the topics from the test. Such interaction can ultimately contribute to participants' achieving better results on the posttest. In order to control for that variable, the researcher clearly explained to the participants that the results were confidential, and that these results would not affect their results in the course they were taking on any possible level. That might have diminished participants' anxiety and concern, ensuring that the results were more genuine in relation to what they have learned due to the treatment alone.

Another issue within *Testing* effects is that of reactivity, which implies that the mere act of measuring something might change what is being measured. In other words, the test itself becomes a direct stimulus for change, and after the event of the pretest, for instance, participants could be compelled to study more during the three-week grammar module in order to improve their performance in the posttest. According to Stanley and Campbell (1963), "the more novel

and motivating the test device, the more reactive one can expect it to be" (p. 9). In order to control for that variable, although the test was not the first assessment of these participants' course, it was the first one related to grammar and syntax. It was administered during the regular class period for thirty minutes only, implying that it would not require extensive preparation within the course routine or generate anxiety for the participants. Besides that, another way to control for *Testing* effects was the fact that the test awarded no points for *performance* (only participation), which also aimed at disconnecting the act of taking part in the study with the need to study for better performances in the tests.

Unreliability of Measure. In relation to a threat to Statistical Conclusion Validity, and the validity aspect of the test used as the Instrument, it was considered valid since it has been used and adapted multiple times and from multiple reliable sources along decades. Nevertheless, a decision to run new analyses for the overall results of the ALAT, including the newly added items, as well as for each section was made in order to strengthen its validity. Results showed an *overall* reliability index of .897, indicating a high degree of consistency for a test of 48 items (see Table 1 in Instrumentation). For the second part of the test (S1P2), because it has only four items, the possibility for guessing is high, which could make this particular portion more unreliable than the other longer parts. A similar possibility is brought up in a study that used a similar version of the ALAT (Clapham, 2001).

Construct Validity

A threat to construct validity that can be considered for the present study is *Mono-operation Bias*, in which one operationalization of a certain construct population might underrepresent the construct of interest (Shadish et al., 2002). Grammar, which is a very broad

54

and complex construct, can neither be fully quantitatively measured nor judged with certainty from the single administration of one assessment. As much as this instrument has been validated by previous users and had updated Cronbach's alpha analyses conducted for the present study, it does not encapsulate all there is to know about grammar and MA. As a result, the study's findings must be analyzed with caution, avoiding either too positive or negative views.

The *Reactivity to the Experimental Situation* is another type of threat to construct validity that can impact the final results of a study. Participants' responses are not solely echoing the treatment to which they were submitted but also their own perceptions of the study. As a consequence, such perceptions become an intrinsic part of the treatment (Shadish et al. 2002). Although it is difficult to control for such a psychologically-driven threat, a clear explanation of the confidential and anonymous aspects of the current study procedures were conducted in each one of the classes where the study was being done to help control for the normal anxiety that any assessment can generate. Also, the fact that there was no possible penalty to participants due to low scores might have helped reduce this particular threat: regardless of their performance, all participants received a bonus incentive in the form of five points simply for participating in the study.

A final type of construct validity threat that this study controlled for was *Experimenter Expectancies* (Shadish et al., 2002) In order to control for this threat, the principal researcher kept the interaction with participants to a minimum, being present only for the explanation of research and the proctoring of at least two groups, but not teaching any class nor over-explaining or justifying the experiment. The principal researcher also observed the treatment in two of the four different group sessions on each day of the treatment, but without any kind of direct

involvement in the instruction or interaction with the participants. The classes for the four groups that took part in the study took place simultaneously. Because of that, the observations were alternated so that the researcher could be in each group at least once a week. More important, a standardized test was used and there was as second rater to ensure consistency in the rating process.

External Validity

One factor that offers threats to external validity for the present study is *Interaction of the Causal Relationship with Outcomes* (Shadish et al., 2002). This factor is directly related to the generalizability of the cause-effect relationship. In other words, depending on the specific outcome for each participant, the treatment's effectiveness might vary. The treatment is not one and the same for all participants, and the participants are not one and the same, even though they belonged to the same target population. Their ability to understand and benefit from the treatment was not necessarily the same, nor was their past experience regarding grammar instruction. For participants already equipped with good notions of grammar, the topics presented in three weeks might have reinforced what they already knew while reminding them of core concepts. However, for a participant who had no basic notions of grammar the treatment might have added confusion to the restricted knowledge they had. A way to control for this threat was to have a biodata portion that clarified if participants had taken a grammar course before the treatment. However, a more comprehensive bio assessment that includes more information on participants' previous grammar knowledge can control for this threat in a better way for future studies.

The aforementioned threat can also be directly related to *Interaction of Causal Relationships with Settings*, since it brings up the issue of cause-effect relationship holding or not, and the generalizability concern. A way to control for such threat is to both expand the description of the population being investigated as well as to minimize the generalizability of the results to the population more closely related to the study demographics.

Instrumentation

The name of the instrument used in this study is Adapted Language Awareness Test (ALAT), and in order to provide a better understanding of its features and history, a summary of at least four known previous instances in which it was used in different studies was provided here. Although the types of population varied, in all the instances the main objectives of the researchers espoused the goal of the present study: investigate grammar knowledge and MA of multiple populations, such as college students directly pursuing language-related degrees, preservice teachers and teachers, both native and nonnative speakers of English. These four examples of studies that used this same instrument were from universities in the United Kingdom. Therefore, some changes were made to the instrument to ensure that it was more adequate for an American audience (the participants in this study), especially in regards to dialectal aspects.

The ALAT was adapted from a similar instrument used by Bloor (1986), in an article discussing university students and the overall status of grammar teaching in modern classrooms, as he investigated how much students were familiarized with grammatical terminology and other related linguistic issues The results were presented only in descriptive statistics, followed by the

author's analyses of the answers and a comparison between the linguists and non-linguists results.

Alderson, Clapham and Steel (1997) adapted the same test from Bloor (1986), as they looked into the state of language knowledge of 509 participants who were also learners of French, most incoming undergraduate students in the United Kingdom. Since communicative goals seemed to make grammar teaching less relevant, the researchers believed that consequently students didn't know as much about grammar as they should. Alderson et al., (1997) were also interested in the notion that lecturers use metalinguistic terminology as they teach grammar topics assuming that students are familiar with it, which is not always true. They also looked at the differences between knowledge *about* language and the functional ability of *using* that same language. This version was a more extensive test battery encompassing multiple sections but it retained parts 1 and 2 of Bloor (1986); other portions were part of a pilot study used by the same researchers.

Andrews (1999b, 2006, 2007) also readapted and used this instrument for investigations of both explicit grammar knowledge and use of grammatical terminology looking specifically at teachers and prospective teachers, as well as looking at native and nonnative aspects of MA and explicit grammar knowledge. A total number of 60 students participated in the study, and the author recognizes that given the size of the groups the ability to generalize its results is limited. Andrews (1999b) reported the results with mean scores and standard deviation for the overall test and the figures for each section of the test. Mean scores on the four measures (the overall test as well as the three sections) were also subjected to a one-way ANOVA, which confirmed the significant variance among the four different groups. This same instrument contained another

task added to the original first section of Bloor (1986) but that section was neither used in Alderson et al., (1997) nor was it used for the present study. Finally, Alderson and Hudson (2012) looked at issues of metalinguistic knowledge and grammar in another study that used test-based survey that gives credit to both Bloor (1986) and Alderson et al., (1997). The study's participants reached 726, with raw scores reported and also a thorough comparison between the previous tests scores, dating from 1986 and the current ones, of 2012.

The aforementioned authors Charles Alderson and Stephen Andrews were contacted by the researcher and provided written authorization for their adapted versions to be used freely (see Appendix C). The authors requested that their previous research studies, in which the tests were used, be cited in the present study.

ALAT The adapted version for the current study

In order to clarify the use of the current instrument as an *adapted* version of the previous ones, for the purpose of this study it was called Adapted Language Awareness Test (ALAT) (Appendix B). In order to preserve the validity of the original instrument, the structure of the original test has remained essentially the same in the adapted version (ALAT), with justified changes made to its previous versions. The ALAT, as it is used in this study, was comprised of two different grammar sections, namely Section 1 and Section 2. Each Section had two parts, Part 1 and Part 2. In order to make it easier to identify the sections and parts, they were called: S1P1, S1P2, S2P1, and S2P2 throughout the study, in which *S* stands for *sections*, and *P* stands for *parts*. (see Appendix B for the ALAT in its entirety)

Section 1. Section 1 contained two separate parts: (a) in Part 1 a single sentence was given, from which participants had to extract 14 different grammatical items, which were listed

for them to choose from. They were: verb, noun, countable noun, passive verb, adjective, adverb, definite article, indefinite article, preposition, relative pronoun, auxiliary verb, past participle, conjunction, finite verb, and infinite verb; (b) in Part 2 four sentences were given, from which participants had to identify four parts of sentence, one for each sentence: subject, predicate, direct object, and indirect object.

Section 2. Section 2 also contained two separate parts: (a) Part 1 encompassed a list of 15 English sentences. Each sentence contained a grammar mistake, which was neither underlined nor in bold. Participants had to recognize the mistakes on their own, underlining the specific faulty part in the sentence and rewriting the faulty part correctly; (b) in Part 2 participants were asked to *explain* the grammatical rule they had previously identified, stating what rule they thought had been broken underneath each sentence. Part 2 of Section 2 was an open-ended format, requesting grammar explanations for the grammatical issue with each sentence and why, and showing participants' level of metalanguage awareness and their knowledge of the grammatical terminology. This part was significantly more challenging in nature than the previous ones and it was expected that participants would either take longer to do it or simply fail to do it altogether, as the pilot test administered on this portion demonstrated in the fall of 2014.

The changes made to the instrument. Although the adaptations made to the original test were not major, they were deemed relevant to ensure the validity of the test and that the items being tested were coherent with the given curriculum for the grammar module. It is important to note that most versions of the test used prior to the current study were in a British setting, with

some terminology choices that were not common in American English, which proved to be distractors in the pilot test. The most relevant adaptations that were made to the instrument were:

Section 1, Part 1. In this portion of the first section of the instrument, *definite article* was added to the list of 14 grammatical items that should be identified. The reason for this addition is that an *article* can be either *definite* or *indefinite*. In the original instrument, as used by Bloor (1986), only an *indefinite article* is required to be identified, when in reality the sentence from which the items must be identified also contains an example of a *definite article*. Alderson et al., (1997), on the other hand, added *definite article* to their version. Because of that, the item *definite article* was added to the list of 14 grammatical items to be identified by participants. Although articles can sometimes be seen as individual parts of speech in some grammar books, in reality the items *the*, *a*, and *an* represent a subset of the part of speech called *adjectives,* not individual parts of speech called *articles* (Folse, 2009). Another change in Part 1 is that the grammatical items *passive verb* was replaced by an example of *uncountable noun*.

Section 1, Part 2. No changes were made for Section 1, Part 2.

Section 2, Parts 1 and 2. In this portion of the second section of the instrument some lexical items were replaced due to semantic distraction concerns that were brought up during the pilot test in the fall of 2014. Participants at the time pointed out to words such as *flat, spanner, colic* and *phoned*, which caused them to either misunderstand the sentence itself, missing the real grammatical issue with the sentence, or simply correct the sentences by 'fixing' these words, as if they were the problematic issues with the sentences. It was hypothesized that these were mere dialectal differences between words more common in British than in American English.

Reliability: Cronbach's alpha. Although previous reports indicated a Cronbach's alpha of .86 for the first portion of the original instrument (Clapham, 2001), new analyses were conducted for the overall results of the ALAT, including the newly added items, as well as for each section. Table 1 depicts measures of internal consistency for the ALAT composite score, Section 1 Part 1, Section 1 Part 2, Section 2 Part 1, and Section 2 Part 2. Results showed an *overall* coefficient of .897, which is considered fairly high for a test of 48 items.

Table 1: Cronbach's Alpha for the ALAT

ALAT Parts	Topic	Number of Items	Points	Cronbach's α
ALAT overall		48	0, 1, 2	.897
Section 1 Part 1	Parts of Speech	14	0, -, 2	.829
Section 1 Part 2	Parts of Sentence	4	0, -, 2	.171
Section 2 Part 1	Correction of Errors	15	0, 1, -	.505
Section 2 Part 2	Explanation of Errors	15	0, 1, 2	.863

ALAT – Mark Scheme

According to Gamaroff (2000), an analysis of the rating processes should also generate a final product that can be used as a guide, in the form of a protocol or script. A thorough analysis of the mark schemes previously used by the aforementioned authors (Andrews, 1999b, 2007; Alderson, Clapham and Steel, 1997), helped with the development of a newer version, containing updated portions with analyses of the new items that were added to Section 2, Parts 1 and 2 of the ALAT (Appendix D).

This new mark scheme was created based on (a) the previous authors' own mark schemes with comments on what should be considered right, wrong or incomplete regarding the answers

given by participants; (b) the principal researcher's own notes after rating both the tests given in the pilot as well as the 202 pre- and post ALAT tests for the current study; and finally (c) the final adjustments that were considered necessary after subsets of samples of the main data set were rated by the second rater as calibration efforts, before the rating of the final 20% subset. The new mark scheme was essential to ensure interrater reliability (Davies et al., 1999), or the level of consensus between the two raters. These are the guidelines of the rating of the ALAT:

Section 1. Section 1 of the ALAT, which includes Parts 1 and 2, was rated very objectively: questions are either right or wrong. For Section 1, Part 1, answers can be considered *correct*, *wrong*, and *not answered*, with two points being assigned for a correct answer, zero point for a wrong answer and zero point for a not answered answer. Section 1, Part 2 contains only four questions, for which there are only four options to be distributed among the four questions; it works like a 'match the columns' type of exercise. The same point distribution is used for Part 2: two points are assigned for a correct answer, zero point for a wrong answer and zero point for a not answered answer.

Section 2. This section had both (a) an error identification part, and (b) an open-ended essay part, where participants explained what was grammatically wrong with each sentence and why. For Part 1, the options are *correct*, *wrong* and *not answered*. Two points were assignedfor *correct*, zero point for *wrong* and zero point for *not answered*. Part 2 rating options are: *fully correct*, *partially correct*, *incorrect*, and *not answered*. A summarized view of the ALAT is depicted on Table 2, showing the two sections (Sections 1 and 2) with their respective parts (Parts 1 and 2), topics being rated , number of items, rating values and total points possible per section, part, and total.

Table 2: ALAT Point Distribution

ALAT Parts	Topic	Number of Items	*Rating Values	Total Points
ALAT overall	All sections	48		81
Section 1 Part 1	Parts of Speech	14	0, -, 2	28
Section 1 Part 2	Parts of Sentence	4	0, -, 2	08
Section 1 Total				36
Section 2 Part 1	Correction of Errors	15	0, 1, -	15
Section 2 Part 2	Explanation of Errors	15	0, 1, 2	30
Section 2 Total				45

Note *0= wrong; 1point=fully correct for Part 1, partially correct for Part 2; 2points= fully correct.

Interrater Reliability. Altogether, 9,696 items were rated in the present study (total for Pretest 4,848; total for Posttest 4,848). Research question 1 encompassed the total number of items; Research Question 2 had a total of 2,828; Research Question 3 had 808 items rated; Research Question 4 had 3,030 items rated; and finally Research Question 5 had 3,030 open-ended analyses rated. A problematic issue in any essay-like language assessment is that the scores are based on raters' judgment of grammatical accuracy, which can vary substantially from one rater to another. According to Gamaroff (2000), rater reliability is more notably relevant in these types of subjective tests "where there exist fluctuations in judgments" (p. 32), both between two different raters (interrater reliability), as well as within the same rater's judgment (intrarater reliability). This is particularly true in relation to the last portion of the ALAT, Section 2, Part 2, which asked participants to name and explain grammatical errors present in 15 sentences. In

extensive analyses of answers from the pilot test, it was clear that this particular portion of the test would need a very accurate mark scheme to ensure consistent ratings. Gamaroff (2000) suggests that using a second rater might control for this kind of problem.

Because of that, a second rater evaluated a randomly chosen 20% of the assessments. The second rater holds a Master's degree in TESOL as well as a PhD in Education with a specialization in TESOL and extensive teaching experience in the area of applied linguistics. A Pearson product-moment correlation coefficient was used to compare the composite scores from the overall scores of the second rater's to the scores given to the same participants by the main rater. Results indicated a positive correlation, $r = 988$, $n = 20$, $p < .001$, demonstrating that there was a high degree of consistency between the two raters in relation to their scores of a 20% subset of the ALAT data set.

However, it is recommended that in future investigations, besides the analysis of the overall scores, separate analyses of each section be also applied. The last portion of the ALAT is the most challenging in terms of achieving objective ratings. In calibration efforts of two samples of 10% done prior to the final 20%, it was clear that the difficulty in achieving higher degrees of consistency in ratings was almost solely concentrated on the final portion, Section 2, Part 2.

Although it is not realistic to expect *full agreement* in subjective answers among raters at all times, it is very important to ensure that the raters have a high degree of consistency overall in the test while assessing the test's reliability. Because of that, measuring the two raters' using a correlation coefficient ensured interrater reliability, and also suggested that the mark scheme is a solid, reliable tool in rating the instrument (Alderson, Clapham, & Wall, 1995).

Data Collection Procedures

   The study was comprised of three phases: (a) the administration of the ALAT as a pretest, along with the bio-data questionnaire; (b) the treatment (a three-week grammar module), comprised of chapters and presentations on morphology, grammar, and syntax, and finally (c) the administration of the ALAT as a posttest to be conducted immediately after the three-week module. Permission to approach the students and explain the research study was given as per IRB process (Appendix A). The instructors of the four sessions of the applied linguistics course had already been approached in December 2014 and agreed to participate in the study as well. Since it was determined that the study did not require signed consents, an Explanation of Research was created to give participants access to the most relevant information about the study. That document was sent to all participants through their university email, and copies were available on the day of the pretest.

   Specifics about the Applied Linguistics Course. The state-mandated undergraduate course lasts for an average of 16weeks in any given spring or fall term, and focuses primarily on issues in second language acquisition. This is one of two mandatory courses that preservice teachers must take to be considered ESOL-endorsed at this urban university. The course is comprised of 15 chapters, in the sequence as follows: language and the brain, first language acquisition, second language acquisition, phonetics, phonology, etymology, morphology, grammar, syntax, semantics, pragmatics, discourse analysis, regional variation, social variation and culture. Typically, three weeks are used to cover the morphology, grammar, and syntax and chapters (in that specific order), with individual variations depending on the instructor, classroom conditions, and number of students, among others. The courses are typically taught by

Master's or PhD students acting as Graduate Teaching Assistants. Online courses are also offered, but were not used for the purpose of this study.

Specifics about ALAT Administration. The ALAT tests were administered during regular class time for a total of 30 minutes only, out of the regular 75-minute class. Participants filled a bio-data questionnaire within the pretest obtaining information about their background, such as *gender, age, major, nationality* and *language* background, as in if they speak *English*, *Spanish* and/or *Other Languages*. Participants were informed that none of this information would negatively affect their grades for the course in which they were registered neither would it be used in any way other than to investigate the issue of knowledge about language and MA. However, five points were granted as an incentive for participants who agreed to take part in the study. As per IRB determination, another option to receive the same number of points was made available for those who declined to or could not participate, so that all students in the class had the opportunity to obtain the same number of credits. All the practical steps and the timeline proposed in this study were explained in detail to the future instructors of record of the applied linguistics course during a meeting in December 2014, prior to the beginning of the spring semester.

Timeline for the investigation. Classes for the spring semester of 2015 started in January on the main campus where this course takes place. The explanation of the research study was done one class day before the first data collection happened. During that class, the principal researcher introduced herself, explained the study and clarified questions from participants. The principal researcher made the IRB form *Explanation of Research* available for all participants; the same form was also sent through registered institutional email to all participants. Since the

study was under *Exempt Status*, there was no need for participants to sign an *Informed Consent* to take part on the study. This presentation took place in each one of the four classrooms with the instructors' permission, and did not take more than 15 minutes of the regular class time.

Around week seven of the course, on February 26th, and immediately before the first grammar component of the course (morphology), participants were administered the bio-data questionnaire and the ALAT pretest for the first 30 minutes of a regular 75-minute class. A script was followed by the person proctoring the pretest, with step by step procedures for administration of the pretest (see Appendix E). Since the four groups had classes simultaneously, the principal researcher was able to proctor only two sessions, together with the instructors of record. The second rater proctored one pretest and one posttest. The person responsible for proctoring the test read from a script that reinforced the main guidelines for participating in the study as well as established the policies for the duration of the test. The IRB form Explanation of Research containing all the information about the study was again made available for all participants in case they had not read it prior to the pretest day. During the remainder of the class, and after the pretest took place, the instructors introduced the first chapter of the morphology component. Participants then participated in the treatment (the three-week grammar module), and immediately after that they took the posttest, on March 31st. The procedures for the posttest were similar to the pretest.

Language Module

Considering the relevance of the treatment in this study (the grammar module), and in order to ensure that replication would occur without missing relevant information, it was

essential to clearly define the three-week language module and contextualize its format., that way making sure that what was given in the module related directly to what tested in the ALAT.

Cross-referencing the ALAT and the Language Module. In order to provide a very clear idea of the intertwined topics that were present in the language module as well as tested in the ALAT, three tables with a cross-reference of the module and test topics were created and are provided below. The tables also contain information about specific assignments that were requested of participants throughout the course, and which of these assignments contained topics that were also present in both the grammar module as well as in the ALAT. In these cases, participants were exposed to these topics during the regular lectures and also prepared specific work related to them in the form of their assignments.

It could be argued by analyzing these tables that since some topics required more direct instruction, participants could show better results in relation to them. It is relevant to reiterate the fact that Research Question 1 addressed the overall increase of scores from pre- to posttests, whereas Research Question 2 to 5 addressed each individual portion of the ALAT concerning difference in scores or lack thereof. Table 3 contains items from Research Questions 1, 2, 4, and 5.

Morphology. The first content covered in the treatment. On average, two classes were dedicated to this chapter.

Table 3: Cross-reference of the ALAT and Morphology

Chapter 6 – Morphology	ALAT
PowerPoint presentation: Morphology (34 PowerPoint slides)	
Definition of morphology;	
Free and bound morphemes: definitions, examples, and examples for analyses;	**Section 2 Parts 1 & 2**
Lexical morphemes: definition, multiple examples;	Affixation processes:
Functional morphemes: definition, multiple examples;	- quick versus quickly;
Derivational morphemes: definition, multiple examples, exceptions;	- taller versus tallest;
Inflectional morphemes: definition, multiple examples, exceptions;	- inflectional morphemes;
The 8 Inflectional morphemes in detail;	- *goed
Morphs, allomorphs, zero morph: definitions and examples;	
Past tense formation analysis: the *ed*;	
Other languages examples	
Morphology-related assignments for the course	**Section 1 Part 1**
Assignment 4: Prescriptive and Descriptive Language	**Section 2 Parts 1 & 2**
Assignment 6: Identifying Errors	
Assignment 7 : Parts of Speech	

Note *ungrammatical example

Grammar. The second content covered in the treatment. Three classes were dedicated to this chapter. Table 4 contains items from Research Questions 1, 2, 4, and 5. This is the most complex portion of the module in relation to the variety of topics that were discussed.

Table 4: Cross-reference of the ALAT and Grammar

Chapter 7 – Grammar and Phrasal Verbs	ALAT
PowerPoint presentation: Parts of Speech and Phrasal Verbs (60 slides) Prescriptive approach/ Descriptive approach: rules and discussion on prescriptive versus descriptive: multiple examples of prescriptive versus descriptive grammar; Infinitive verb; Parts of speech: Nouns, Verbs, Pronouns, Adjectives, Adverbs, Conjunctions, Articles, Prepositions, Interjections: slides defining and exemplifying parts of speech one by one; multiple examples analyses; Verb Tenses: examples and discussion; verb tense issues for ELLs; Pronouns and apostrophes issues: multiple examples and analyses; Grammatical gender: multiple examples and analyses; Agreement issues: multiple examples and analyses; Phrasal verbs: definitions, examples, separable and inseparable PVs, phrasal verbs and phrasal nouns. **Extra PPT on Grammar**: Pronouns, Adjectives and Prepositions (18 slides) Pronouns: applicability, types of pronouns, examples, analyses, and practice; Adjectives: applicability and examples; Prepositions: applicability, examples, analyses, exceptions and comparison between British and American usage. **Extra PPT on Grammar**: Grammar Supplements (10 slides) Pronouns; Verbs: Focus on verb To be and Contractions with To be;	**Section 1 Part 1** -14 items including 7 parts of speech, + infinitive verb **Section 2 Parts 1 & 2** -subject-verb agreement; -pronoun use; -preposition; -descriptive versus prescriptive grammar; -double negative; -use of phrasal verbs: separable and not separable; -verb tenses and verb tense formation and use; -affixation processes

Grammar-related assignments for the course

Assignment 4: Prescriptive and Descriptive Language	**Section 1 Part 1**
Assignment 5: Phrasal Verbs	**Section 2 Parts 1& 2**
Assignment 6: Identifying Errors	

Syntax. The third content covered in the treatment. One or two classes were dedicated to this chapter. Table 5 below contains items present in Research Questions 1 and 3.

Table 5: Cross-reference of ALAT and Syntax

Chapter 8 - Syntax	ALAT
PowerPoint presentation: Parts of Sentence and Word Order (26 slides)	
Definition of parts of sentence;	**Section 1 - Part 2**
Subject and predicate;	**-**subject;
Simple, complete, and compound subject: examples;	-predicate;
Simple and compound predicate: examples;	-direct object;
Objects: direct and indirect: definitions, examples, analyses, practice;	-indirect object
Word order in English: SVO;	
Word order in other languages: examples	
Syntax-related assignments for the course	**Section 2 – Parts 1, 2**
Assignment 6: Identifying Errors	

GTAs: An Important Aspect of the Language Module. In general, the four face-to-face sections of this course are taught by graduate teacher assistants, or GTAs, who are either

Master's or PhD students/candidates and teach this course as part of their assistantships. Because of that, GTAs do not usually teach this course for more than one or two terms, with few exceptions.

The teaching experience as a GTA is expected to be an autonomous one that allows each instructor to develop his or her own individual approaches to teaching. That means that there is no clearly defined script by which the course content is taught. For instance, some GTAs might choose to follow the book chapters more strictly and follow PowerPoint presentations slide by slide, while others approach the classes from a more hands-on perspective and use in-class activities, games, or debates, depending on the topic at hand. In the same vein, and as stated before, some GTAs might prioritize topics in the chapters that are also present in the assignments for the course, in a way catering to both content knowledge as well as a direct support to the proposed coursework.

A great number of PowerPoint presentations corresponding to all the book chapters was available for the GTAs, including the ones for grammar, morphology and syntax. However, instructors were free to use them as they were, add materials to them, or prepare their own presentations or activities. Because of time constraints, GTAs who have taught this course have usually opted for the available ready-to-use material, with no or few changes.

Meetings prior to the beginning of the spring semester between the researcher and the GTAs ensured that the GTAs were aware of the research main guidelines and that they agreed with the proposed calendar. Both GTAs of record for the four groups were instructed not to divert from the basic grammar, syntax and morphology chapters available from the course's presentations in order to avoid any extra amount of instruction that could interfere with the final

results and give unfair advantage to one group over another. Nevertheless, stated, there was freedom to teach the classes in multiple ways, depending on each GTA's perspectives on each topic.

For the present study, both GTAs teaching the courses were TESOL Master's students; one had taught the course once before in the fall of 2014, while the other was teaching this course for the first time in the spring of 2015.

Observing the Language Module. Since the individual characteristics of each GTA while teaching grammar components were not the same, and to control for external validity, or more specifically, *interaction of the causal relationship over treatment variations* (Shadish et al., 2002), the principal researcher observed as many sessions as possible during the three-week module, and made individual notes of what was present in the chapters and what was taught by the instructors and how. By doing so, the principal researcher ensured a higher level of accuracy in relation to the correspondence between course content, or what was covered during the language module and the ALAT.

Neither *attendance* nor *instructor* was part of the primary research questions of the present study. However, a record of participants' attendance was also kept in order to control for that variable for supplemental analyses. The same can be said about the instructors. Although investigating the difference in the way each instructor taught the classes was not a primary concern of the present study, this factor was analyzed as a supplemental analysis. The principal researcher, who used to be an instructor for the same course, was not involved with any aspect of teaching the classes, being present only for the purposes of explaining the research, administering the pre- and posttests, and observing the classes. That allowed for a more objective

and removed perspective on the results, also controlling for construct validity, more specifically

experimenter expectancies (Shadish et al., 2002).

<center>**CHAPTER 4:**</center>

<center>**RESULTS**</center>

The present study employed an initial set of five statistical analyses (paired-samples *t*-tests) in order to investigate the five primary research questions. The independent variable was the treatment, in this case the three-week module in which participants took part during one of their courses. The dependent variable, which measured the effect of the treatment or lack thereof, was the MA of participants in this study, or more specifically the scores obtained in two tests (pre- and post-) that measured their MA (Lomax, 2007). The dependent variables were investigated directly in relation to each of the five research questions for the present study.

Summary of Results

Paired-samples *t* tests were used to investigate the research questions in relation to scores in the pretest and posttest (after the *Treatment)* using the whole dataset (*N*=101). Subsequently, the tests were rerun after the removal of 11 participants who did not attend at least five classes out of the total number of six classes for the whole treatment (*N*=90). According to principles recommended by Green and Salkind (2008), effect sizes *d* were hand-calculated using the following formula (p. 165):

$$d = \frac{t}{\sqrt{N}}$$

The paired-samples *t*-tests for Research Questions 1, 3, 4, and 5 were also rerun after the detection of outliers. No outliers were detected for Research Question 2. A series of investigations were performed to determine normality, and Shapiro-Wilk results are available under each Research Question. Despite the acknowledgment that an assumption of normality

<center>76</center>

was violated, the analyses of the *t*-tests results showed consistency between before and after the outliers removal, with the *t*-tests still being statistically significant for Research Questions 1,3, and 4, and not statistically significant for Research Question 5. More details regarding the removal of outliers will be available under each Research Question results, as well as under the Discussion chapter.

Throughout the Results Chapter, Research Question 1 will always be referred to as *overall results*. Research Questions 2, 3, 4, and 5 will always be referred to by the abbreviations of their respective sections/parts of the ALAT: *S1P1, S1P2, S2P1*, and *S2P2*.

Table 6: Descriptive Statistics for Research Questions 1, 2, 3, 4, and 5

	Maximum Points		Mean	Standard Deviation	Standard Error Mean	N
Res. Question 1 (Overall scores)	81	Pre	34.386	13.3566	1.3290	101
		Post	40.901	13.1244	1.3059	
Res. Question 2 (S1P1)	28	Pre	11.980	7.0965	.7061	101
		Post	15.931	6.8414	.6807	
Res. Question 3 (S1P2)	8	Pre	3.832	1.8389	.1830	101
		Post	5.644	1.9266	.1917	
Res. Question 4 (S2P1)	15	Pre	12.545	1.7974	.1788	101
		Post	12.891	1.6787	.1670	
Res. Question 5 (S2P2)	30	Pre	6.030	5.3413	.5315	101
		Post	6.436	5.6788	.5651	

Statistical Results

Research question 1. Is there a statistically significant increase in the overall metalinguistic awareness of participants after a treatment, as measured by pre- and post-administrations of the Adapted Language Awareness Test (ALAT)?

Research hypothesis 1. The directional hypothesis stated that there is a statistically significant increase (p < .05) in the overall metalinguistic awareness of participants after a treatment, as measured by pre- and post-administrations of the ALAT. The null hypothesis (H0: $\mu 1 - \mu 2 = 0$) was rejected in favor of an overall increase in performance (Ha: $\mu 1 - \mu 2 > 0$).

Result 1. In addressing Research Question 1, a paired-samples *t*-test was conducted to determine whether there was a statistically significant increase in relation to overall scores from pre- to posttest of the ALAT. Results indicated that the mean of the pretest (*N*=101, *M*= 34.38, *SD*= 13.35) had a statistically significant increase from pre- to posttest (*M*= 40.90, *SD*= 13.12, *t*(100)= .9.10, *p*<.001. The 95% confidence interval for the mean difference ranged from -7.9350 to -5.0947. The effect size index *d* was .90, which is considered a large effect size under the conventional interpretation of .2, .5, and .8, representing small, medium, and large effect sizes, respectively (Green & Salkind, 2008). In addition, Figure 1 below depicts the mean scores of each part of Research Question1.

In Figure 1, the horizontal axis represents the four individual parts of the ALAT: Section 1, Part 1 (S1P1), Section 1, Part 2 (S1P2), Section 2 Part 1 (S2P1), and finally Section 2, Part 2 (S2P2). Each part is related to Research Questions 2, 3, 4, and 5, and their respective parts of the ALAT. Research Question 2 is represented by *Part 1*; Research Question 3 is represented by *Part 2*; Research Question 4 is represented by *Part 3*; and finally Research Question 5 is

represented by *Part 4*. In Figure 1, the vertical axis represents the average percentage of participants' scores (from 0 to 100%) of the four pre- and posttests of each part that composes the ALAT. Figure 1 offers an initial glance at the heterogeneity of the data, which will be further addressed in more detail in Chapter 5.
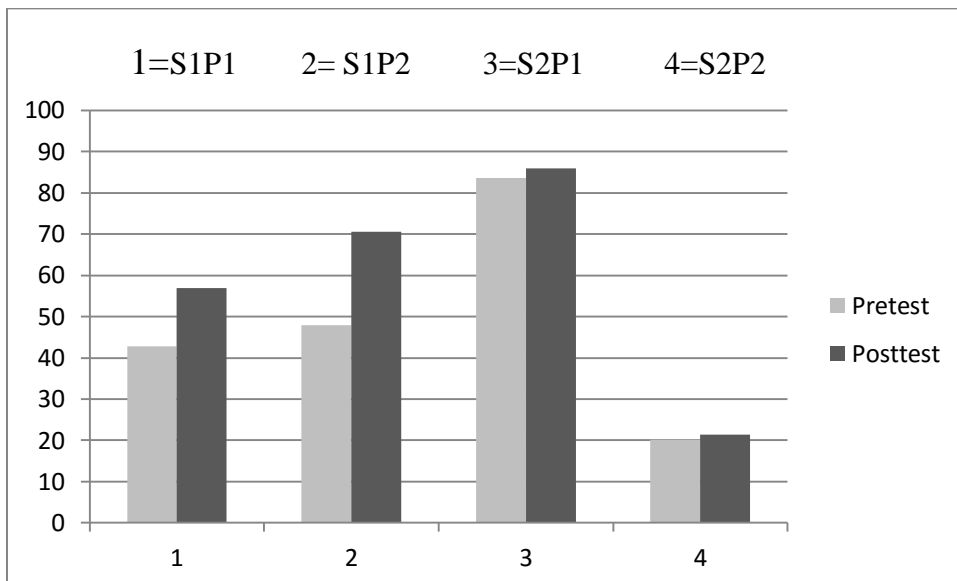
Figure 1: Research Question 1 Overall results of ALAT by Section and Part.
Pretest (light gray bars) Posttest (black bars)

It should be noted that in relation to Research Question 5, represented in figure 1 by 4(S2P2), the rating of the proposed questions was different from the previous parts. This difference and its visible effects on the graph will be addressed under the Results section for Research Question 5.

Figure 2 illustrates the presence of four outliers in boxplots for the pretest and three outliers for the posttest in the analysis of Research Question 1, *Overall results*. The vertical axis in Figure 2 represents the possible scores from 0 to 81 points (where 81 points are the highest

possible score). The horizontal axis represents the pretest (TOPRE) and the posttest (TOPOS).

Subsequent Figure 3 shows the boxplots after the removal of outliers.
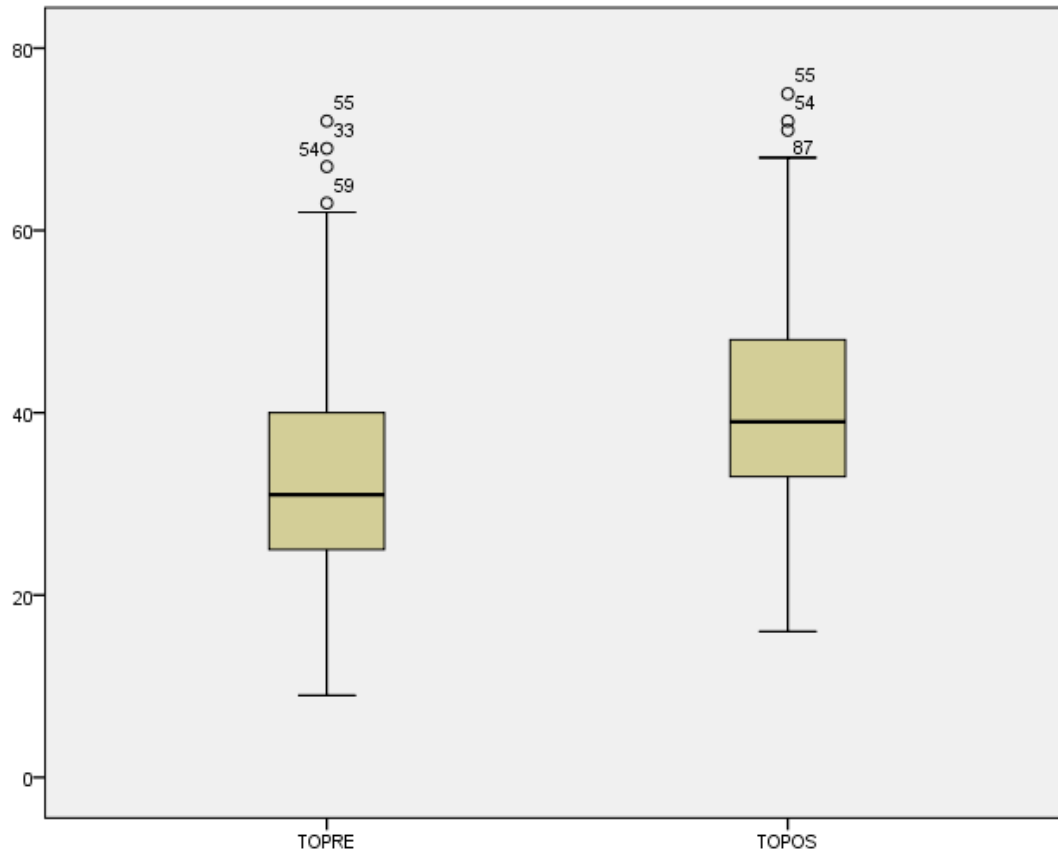
Figure 2: Research Question 1 Outliers of Overall Results

The presence of outliers pointed to evidence that the assumption of normality was violated. In order to account for normality distribution and ensure the tests validity, a series of tests were conducted to account for normality, skewness and kurtosis values. A review of the Shapiro-Wilk's test for normality in the pretest ($W=.941$, $p<.001$) showed statistical significance, which means that it did not meet the assumption for normality. However, skewness (.827) and kurtosis (.234) statistics, which both fell within the acceptable range from -1.0 to +1.0, indicated a reasonable assumption of normality for the pretest. Review of the Shapiro-Wilk's test for

normality in the posttest (W=.971, p=.028) did not meet the assumption for normality. However, skewness (.510) and kurtosis (-.157) statistics indicated that normality is a reasonable assumption for the posttest, since these values were both within the range of -1.0 to +1.0 for skewness and kurtosis. The presence of four high-end outliers in the pretest reinforces the notion of how heterogeneous the data were, with participants scoring from a minimum of 9 (out of 81) to almost maximum points (72 out of 81 points).
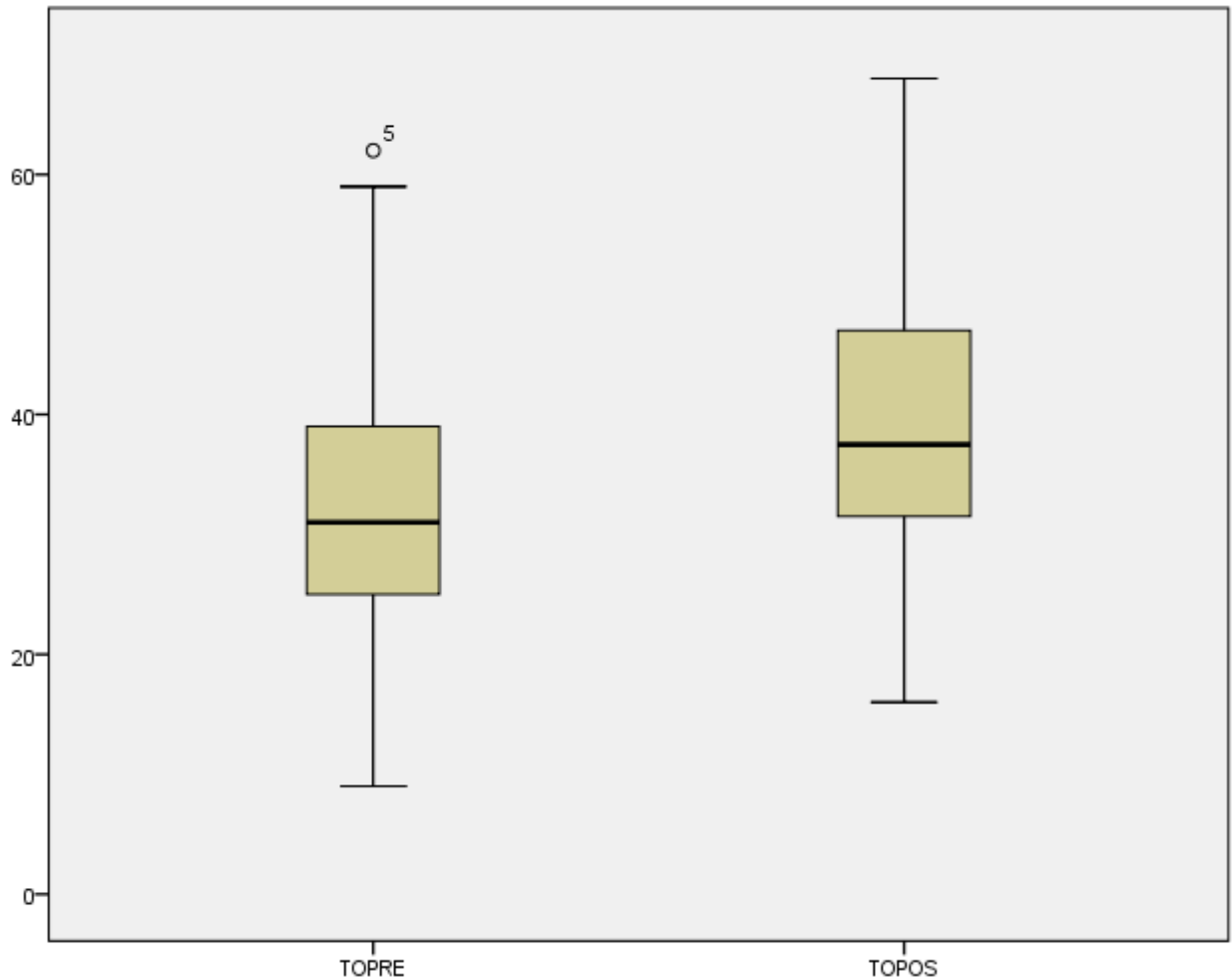
Figure 3: Research Question 1 Boxplot after the removal of outliers

Nevertheless, even after the treatment the range of scores was still compelling in regards to the data heterogeneity. Scores started at 16 and reached up to 75, with high-end outliers again at the maximum scores. A decision to remove the outliers was made in order to ensure that the tests were still robust despite the outliers' detection. In effect, *t*-tests were rerun and the results continued to show statistically significant increase in overall scores from pre- (*M*= 32.76, *SD*= 11.50) to posttest (*M*= 39.406, *SD*= 11.61), *t*(95)= 8.98, *p*<.001, suggesting that the test remained robust despite the violation of statistical assumption of normality.

There were no outliers present in boxplots for Research Question 2. However, there were outliers present in analyses for Research Questions 3, 4, and 5. Subsequent tests were conducted after removing the outliers for each of the sections where they occurred, but in the same fashion as Research Question 1, there was no difference in the statistical significance of the results as compared to results prior to the removal of outliers for these research questions.

Research question 2. Is there a statistically significant difference in the ability of participants to identify *parts of speech* after a treatment as measured by pre- and post-administrations of Section 1, Part 1 (S1P1) of the ALAT?

Research hypothesis 2. The non-directional hypothesis stated that there is a statistically significant difference ($p < .05$) between the scores of a pre- and post-administration of the ALAT in relation to identification of grammatical items, mostly parts of speech. The null hypothesis (H0: $\mu1 - \mu2 = 0$) was rejected in favor of a difference in performance (Ha: $\mu1 - \mu2 \neq 0$).

Result 2. In addressing Research Question 2 (S1P1), related to the investigation of *parts of speech*, a paired-samples *t* test was conducted to determine whether there was a statistically significant difference in scores from pre- to posttest of the ALAT. The results indicated that there was a statistically significant difference from pre- (*N*=101, *M*= 11.98, *SD*=7.09) to posttest (*M*= 15.93, *SD*=6.84), *t*(100)= 7.87, *p*= .000. The 95% confidence interval for the mean difference ranged from -4.9456 to - 2.9554. The standardized effect size index *d* was .78, which is considered a large effect size under the conventional interpretation of .2, .5, and .8, regardless of sign, representing small, medium, and large effect sizes, respectively (Green & Salkind, 2008).

Research Question 2 was based on the analysis of one sentence, from which 14 parts of speech had to be identified. The sentence was: "Materials are delivered to the factory by a

supplier, who usually has no technical knowledge but who happens to have the right contacts."

The parts of speech, by item number, that had to be identified were as follows: (1) verb, (2) auxiliary verb, (3) participle, (4) infinitive, (5) noun, (6) countable noun, (7) uncountable noun, (8) relative pronoun, (9) adjective, (10) definite article, (11) indefinite article, (12) adverb, (13) preposition and (14) conjunction.

Figure 4 depicts pre- and posttest scores of each of the 14 items assessed in Section 1 Part 1 (S1P1), providing a more detailed view of what was tested in that particular part of the ALAT. For this part, participants identified *parts of speech* from a sentence. The relevance of this figure lies in the fact that it depicts the pre- and posttest participants' percentages, and endorses the inferential findings with descriptive illustration of the item level. The vertical axis in Figure 4 represents the percentage of participants who scored that specific item correctly (from 0 to 100%). The horizontal axis represents each of the 14 items with their pre- and posttest results, which will be explained in further detail.
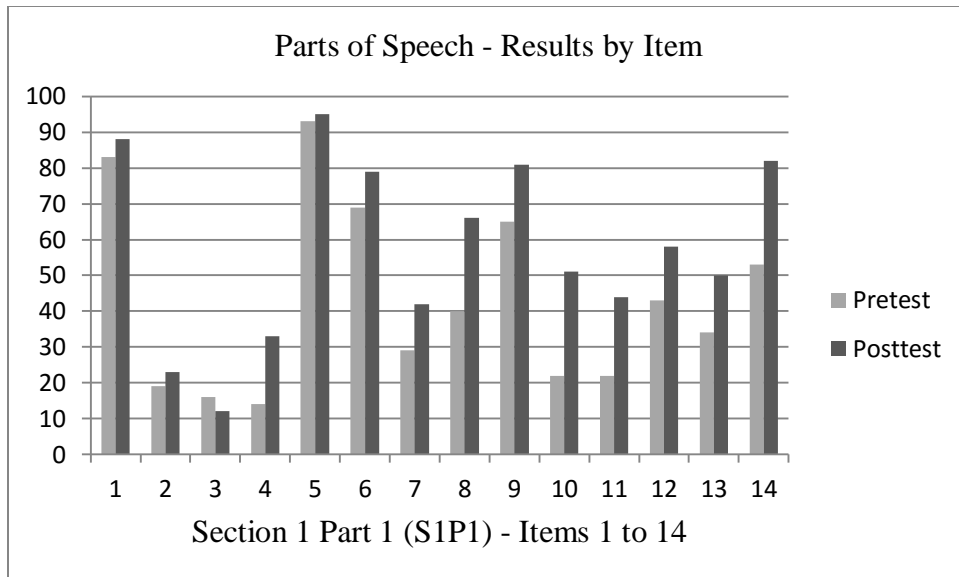
Figure 4: Research Question 2 Section 1 Part 1 (S1P1) Descriptive results by item

Table 7 depicts Research Question 2, Section 1 Part 1 (S1P1), with an itemization of

scores of pre- and posttest for each item in *parts of speech*. In this part of the ALAT, participants

had to identify a number of *parts of speech* from within one given sentence. Table 7

demonstrates in descriptive statistics analyses the percentage of participants who scored

positively for each *part of speech* in the pretest and after the treatment (posttest). Item 5(*noun*),

had the highest percentage of positive scores of all the 14 items, with 93% participants

identifying it correctly in the pretest and 95% in the posttest. The item *verb* had the second

highest percentage of participants' positive scores in both the pretest (83%) and posttest (88%).

An interesting item to be mentioned is Item 3(*participle*), which had a percentage of only 16% of

the participants being able to correctly identify it in the pretest, and the posttest at 12%.

However, Item 4, *infinitive*, had a pretest participants' percentage of positive scores of 14%, and

a posttest percentage of 33%.

Table 7: Research Question 2 Parts of Speech – Itemization of Mean Scores (S1P1)

Item Number	Part of speech	Pretest	Posttest
1	Verb	83%	88%
2	Aux. Verb	19%	23%
3	Participle	16%	12%
4	Infinitive	14%	33%
5	Noun	93%	95%
6	Count. noun	69%	79%
7	Un. noun	29%	42%
8	Rel. pron.	40%	66%
9	Adjective	65%	81%
10	Def. article	22%	51%
11	Ind. article	22%	44%
12	Adverb	43%	58%
13	Preposition	34%	50%
14	Conjunction	53%	82%

Although it is not possible to make inferential claims based on descriptive statistics, and the research questions for the present study were not aiming at item level statistical significance, the descriptive statistical results suggest a need for further investigation in an attempt to provide further evidence that the treatment was indeed effective at each *item level*. It is suggested for future research that McNemar tests for significance of changes (Salkind, 2008) be run to investigate before and after changes at each item level in the ALAT.

Research Question 2 was the only one among the five research questions that presented no outliers in the boxplots, which can indicate that overall participants had a more homogenous

initial stage of knowledge on the items being investigated, and progressed similarly to higher scores from pre- to posttest after the treatment. The interquartile range was also more evenly distributed (10 for the pretest and 8 for the posttest) with a median of 12 in the pretest (out of 28 possible points), and 16 in the posttest. Nevertheless, both pre- and posttest had participants who scored a minimum of 0 and participants who scored a maximum of 28 points.

Research question 3. Is there a statistically significant difference in the ability of participants to identify *parts of sentence* after a treatment as measured by pre- and post-administrations of Section 1, Part 2 (S1P2) of the ALAT?

Research hypothesis 3. On the subject of Research Question 3, the non-directional hypothesis stated that there is a statistically significant difference ($p < .05$) between the scores of a pre- and post-administration of the ALAT in relation to identification of parts of sentence. The null hypothesis (H0: $\mu1 - \mu2 = 0$) was rejected in favor of a difference in performance (Ha: $\mu1 - \mu2 \neq 0$).

Result 3. In order to address Research Question 3, a paired-samples *t* test was conducted to determine whether there was a statistically significant difference in scores from pre- to posttest of the *parts of sentence* on the ALAT. The results indicated a statistically significant difference from pre- ($N= 101$, $M= 3.83$, $SD= 1.83$) to posttest ($M= 5.64$, $SD= 1.92$, $t(100)= 7.604$, $p=.000$). The 95% confidence interval for the mean difference ranged from -2.2846 to – 1.3392. The standardized effect size index *d* was calculated at .75, which is considered a large effect size under the conventional interpretation of .2, .5, and .8, regardless of sign, representing small, medium, and large effect sizes, respectively (Green & Salkind, 2008).

Figure 5 depicts the results of each of the four items of Section 1, Part 2 (S1P2), related to Research Question 3. Participants had to identify four parts of sentence, as follows: (1) subject, (2) predicate, (3) direct object, and (4) indirect object. The four individual sentences were, as follows: (1) Poor little Joe stood out in the snow; (2) Joe has nowhere to shelter; (3) The policeman chased Joe down the street; and (4) The woman gave him some money.

The vertical axis in Figure 5 represents the possible scores as percentages (from 0 to 100%), whereas the horizontal axis represents the four numbered parts of sentence, 1= subject, 2= predicate, 3= direct object, and finally 4=indirect object. It is clear that subject was the item more consistently recognized, with 91% of participants recognizing it in the pretest and 95% in the posttest. A very low percentage of participants correctly recognized the item predicate in the pretest (24%); but 57% of the participants were able to recognize it in the posttest, after the treatment.
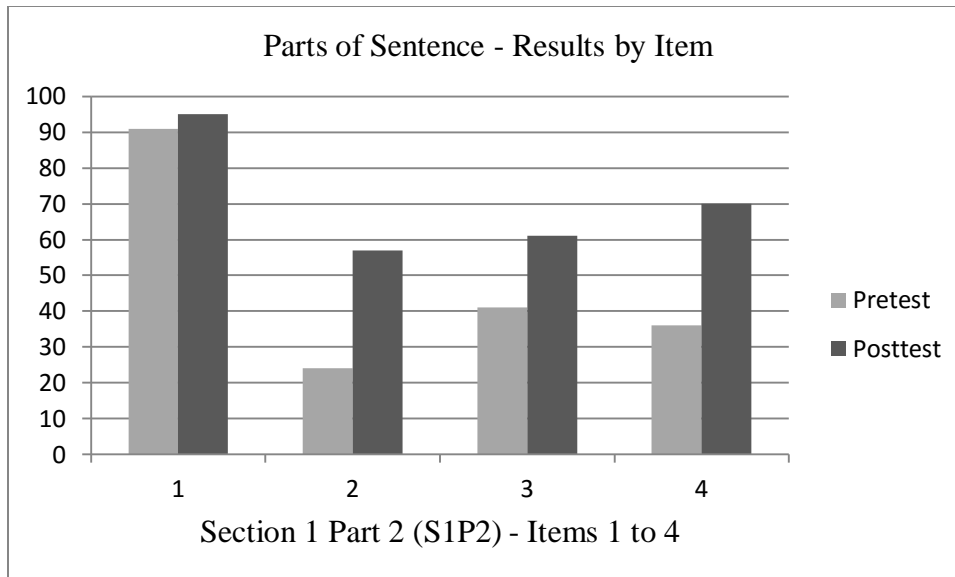
Figure 5: **Research Question 3 Section 1 Part 2 (S1P2) Results by item**

The presence of outliers for Research Question 3 indicated that the assumption of normality was violated. For this reason, and in order to ensure the tests validity, and address normality distribution, a series of tests were conducted to account for normality, skewness, and kurtosis values. A review of the Shapiro-Wilk's test for normality in the pretest (W=.897, p<.001) showed statistical significance, which means that it did not meet the assumption for normality. However, skewness (.224) and kurtosis (.-147) statistics, which both fell within the acceptable range from -1.0 to +1.0, indicated a reasonable assumption of normality for the pretest. Review of the Shapiro-Wilk's test for normality in the posttest (W=.872, p<.001) did not meet the assumption for normality. However, skewness (.-386) and kurtosis (-.501) statistics indicated that normality is a reasonable assumption for the posttest, since these values were both within the range of -1.0 to +1.0 for skewness and kurtosis. Besides these findings, *t*-tests were rerun after the removal of outliers and the results continued to show statistically significant difference in Section 1, Part1 (S1P1) scores from pre- (*M*= 3.66, *SD*= 1.66) to posttest (*M*= 5.54,

$SD=$ 1.90), $t(96)=$ 7.69, $p<.001$, which suggests that the test was robust despite the violation of statistical assumption of normality.

Research question 4. Is there a statistically significant difference in the ability of participants to identify and correct *grammatical errors* in a sentence after a treatment as measured by pre- and post-administrations of Section 2, Part 1 (S2P1) of the ALAT?

Research hypothesis 4. On the subject of Research Question 4, the non-directional hypothesis states that there is a statistically significant difference ($p < .05$) between the scores of a pre- and post-administration of the ALAT in relation to ability to identify and correct grammatical mistakes. The null hypothesis (H0: $\mu1- \mu2= 0$) was rejected in favor of a difference in performance (Ha: $\mu1 - \mu2 \neq0$).

Result 4. In order to address Research Question 4, a paired-samples *t*-test was conducted to determine whether there was a statistically significant difference in scores from pre- to posttest of the *grammatical errors* on the ALAT. The results indicated that the mean of the pretest ($N=$ 101, $M=$12.54, $SD=$ 1.79) was significantly different from pre- to posttest ($M=$ 12.89, $SD=$ 1.67), $t(100)=$ 2.421, $p=.017$. The standardized effect size index d was calculated at .24, which is considered a small effect size under the conventional interpretation of .2, .5, and .8, regardless of sign, representing small, medium, and large effect sizes, respectively (Green & Salkind, 2008).

For this Section 2 Part 1 (S2P1), participants had to identify one grammatical mistake in each of the following 15 sentences: (1) I walk to work very quick; (2) When her said that, Jack hit her; (3) Every day I am making good resolutions; (4) She's the taller of the four sisters;(5) I live in a room at a top of an old house; (6) Mommy goed to the park yesterday; (7) The children put on their coat; (8) He usually like to study at the library;(9) I don't like people which are

always apologizing; (10) I opened the door, but I couldn't see nobody; (11) When I was a small baby I have earaches; (12) I will pick up you later; (13) Josh and Pete have went to the show; (14) Give the paper to Joe and I; and (15) She has called a few minutes ago.

Figure 6, related to Research Question 4, depicts the scores of pre- and posttest for each one the 15 items in Section 2, Part 1 (S2P1). The vertical axis represents the percentage of participants (from 0 to 100%), who were able to accurately identify the mistakes in each sentence. The horizontal axis represents each of the 15 sentences containing grammatical errors.

Most participants were able to accurately identify the mistakes and correct the 15 sentences, with the exceptions being Items 3 and 13. Regarding Item 3 (Every day I am making good resolutions), only 59% of the participants were able to identify and correct the mistake in the pretest, and only 67% of the participants were able to identify and correct the mistake in the posttest. Regarding Item 13(Josh and Pete have went to the show), only 41% of the participants were able to identify and correct the mistake in the pretest, and only 50% of the participants were able to identify and correct the mistake in the posttest.
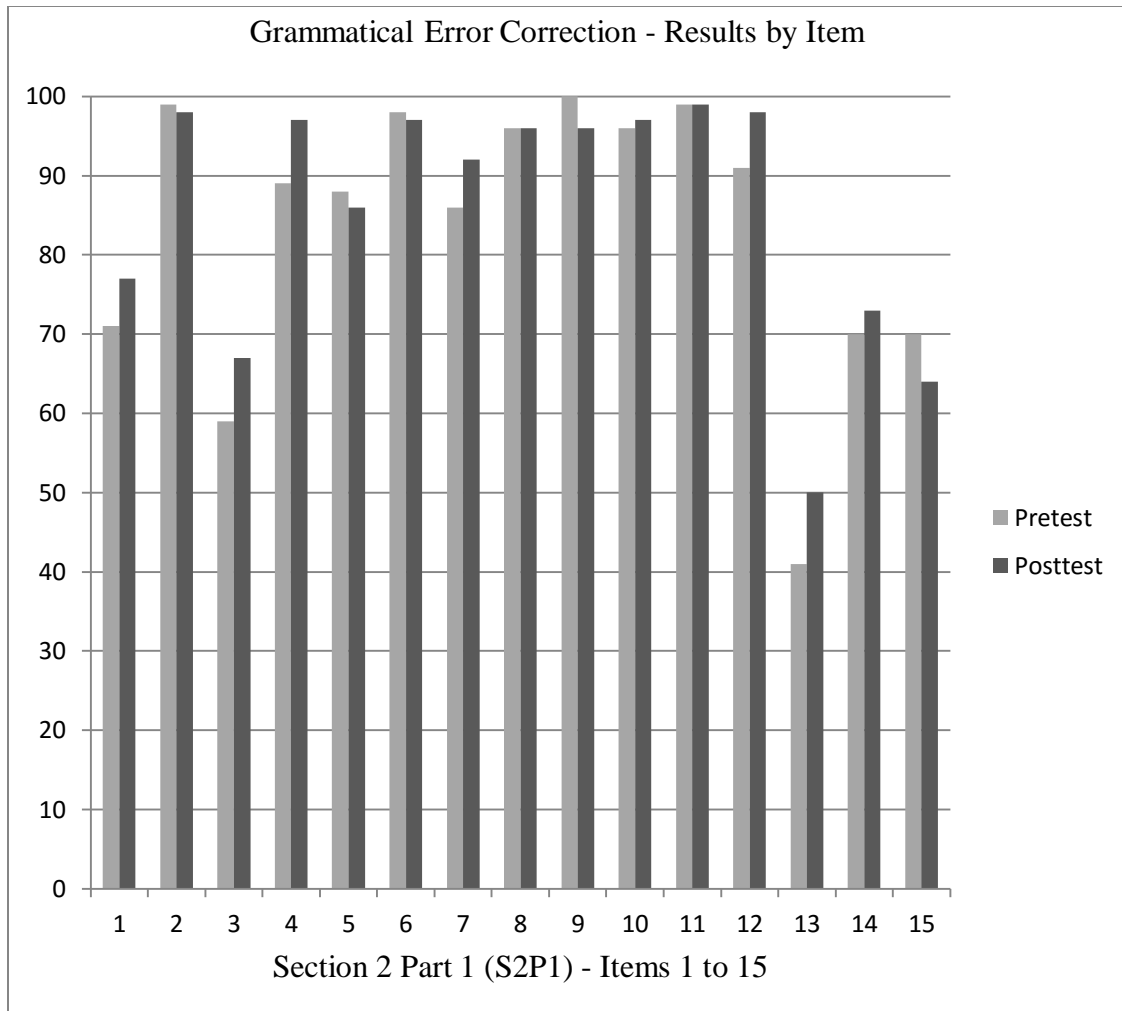
Figure 6: Research Question 4 Section 2 Part 1(S2P1) Results by item

The presence of outliers for Research Question 4 indicated that the statistical assumption of normality was violated. Therefore, and in order to ensure the tests validity, as well as control for normality distribution, a series of tests were conducted to account for normality, skewness and kurtosis values. A review of the Shapiro-Wilk's test for normality in the pretest (W=.882, p<.001) showed statistical significance, which means that it did not meet the assumption for normality. Both skewness (.-1.386) and kurtosis (3.233) statistics fell outside the acceptable range from -1.0 to +1.0, indicating the violation of assumption of normality for the pretest.

Review of the Shapiro-Wilk's test for normality in the posttest ($W=.819$, $p<.001$) did not meet the assumption for normality. Both skewness (-1.960) and kurtosis (6.630) statistics appeared to violate the assumption of normality for the posttest, since these values were both outside the range of -1.0 to +1.0 for skewness and kurtosis. Despite these findings, *t*-tests were rerun after the removal of outliers, with results showing statistically significant difference in scores from pre- (*M*= 12.76, *SD*= 1.41) to posttest (*M*= 13.10, *SD*= 1.22), $t(96)= 2.55$, $p<.012$ in Section 2, Part 1 (S2P1). These findings suggest that despite the violation of statistical assumption the test remained robust.

Research question 5. Is there a statistically significant difference in the ability of participants to provide *metalinguistic explanations for grammatical errors* after a treatment as measured by pre- and post-administrations of Section 2, Part 2 (S2P2) of the ALAT?

Research hypothesis 5. On the subject of Research Questions 5, the non-directional hypothesis stated that there will be a statistically significant difference ($p < .05$) between the scores of a pre- and post-administration of the ALAT in relation to ability to name and explain the same aforementioned grammatical errors. The null hypothesis (H0: $\mu1- \mu2= 0$) failed to be rejected in favor of a difference in performance (Ha: $\mu1 - \mu2 \neq 0$).

Result 5. In order to address Research Question 5, a paired-samples *t* test was conducted to determine whether there was a statistically significant difference in scores from pre- to posttest on *metalinguistic explanations for grammatical errors* of the ALAT. The results indicated no statistically significant difference from pre- (*N*= 101, *M*= 6.030, *SD*= 5.3413) to posttest (*M*=, 6.436, *SD*=5.6788), $t(100)= -1.358$, $p=.177$. The standardized mean difference was 0.135, which is considered a small effect.

It is relevant to note that the rating of Section 2, Part 2 (S2P2), which pertains to Research Question 5, differed from the other parts of the ALAT, since participants could receive a rating of 0, 1, or 2 for each item of that section. Figure 5 below, shows the results of the number of participants who scored 0, 1, or 2 per item number for the pre- and posttest.

Figure 7 depicts the low percentage of overall results for each one of the sentences that had to be corrected as part of the Section 2 Part 2 (S2P2) of the ALAT.
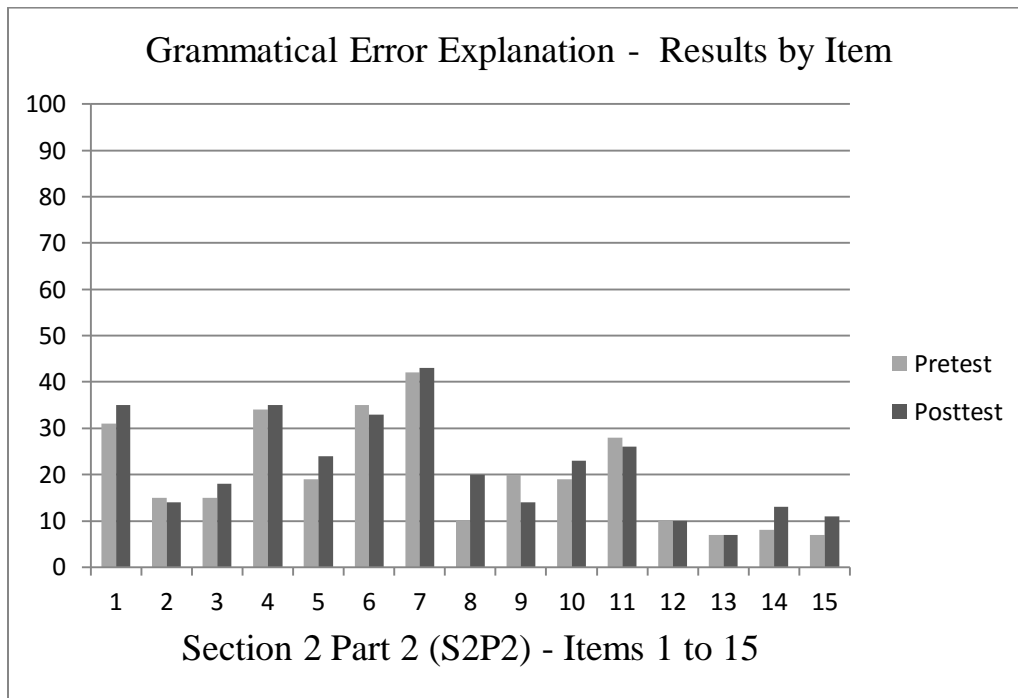
Figure 7: **Research Question 5 Section 2 Part 2 (S2P2) Results by item**

Figure 7 is compelling in illustrating that, not only were participants unable to accurately explain the errors they had just identified successfully in Research Question 4, but in some cases (such as Items 2, 6, 9, and 11), the percentage of participants scoring positively was *lower* on the posttest. These items were: (2) When her said that, Jack hit her, with pre- at 15% and posttest at 14%; (6) Mommy goed to the park yesterday, with pre- at 35% and posttest at 33%; (9) I don't

like people which are always apologizing, with pre- at 20% and posttest at 14%; and finally (11) When I was a small baby I have stomachache, in which 28% of the participants scored positively in the pretest and 26% in the posttest.

Table 8 illustrates the descriptive results by each possible score per item tested. It is interesting to note that the final percentage of participants who scored 0 in both the pre- and the posttest overall scores was very high (68% on the pretest and 67% on the posttest). Even if the results of scores 1 and 2 are combined, they will still be around half of the percentage of 0 score results.

Figure 8 depicts the occurrence of outliers for Research Question 5, Section 2, Part 2 (S2P2), clearly illustrating the heterogeneity of the group of participants in relation to their ability to explain grammatical errors. The vertical axis in Figure 8 depicts the possible scores in points, ranging from zero to 30 points. The horizontal axis represents the total scores for the pre- and posttest. The interquartile range, where most of the participants scored, went from three to eight, with a median of five points (out of 30).

Table 8: Research Question 5 – Section 2, Part 2 (S2P2)	PreTest			Post Test		
	Possible points: 0, 1, 2			Possible points: 0, 1, 2		
Question	0	1	2	0	1	2
	Number of participants who scored each point					
1. I walk to work very quick	59	21	21	52	28	21
2. When her said that, Jack hit her.	76	13	12	83	8	9
3. Every day I am making good resolutions.	8	11	10	75	14	11
4. She's the taller of the four sisters.	52	30	19	48	35	18
5. I live in a room at a top of an old house.	72	19	10	67	19	15
6. Mommy goed to the park yesterday.	42	47	12	45	45	11
7. The children put on their coat.	25	67	9	34	48	19
8. He usually like to study at the library.	88	5	8	68	25	8
9. I don't like people which are always apologizing.	63	36	2	77	20	4
10. I opened the door, but I couldn't see nobody.	73	17	11	63	29	9
11. When I was a small baby I have earaches.	59	28	14	61	28	12
12. I will pick up you later.	85	12	4	85	11	5
13. Josh and Pete have went to the show.	87	13	1	91	6	4
14. Give the paper to Joe and I.	87	11	3	80	15	6
15. She has called a few minutes ago.	90	8	3	83	14	4
Total percentage of participants' scores per each point	68%	22%	10%	67%	23%	10%
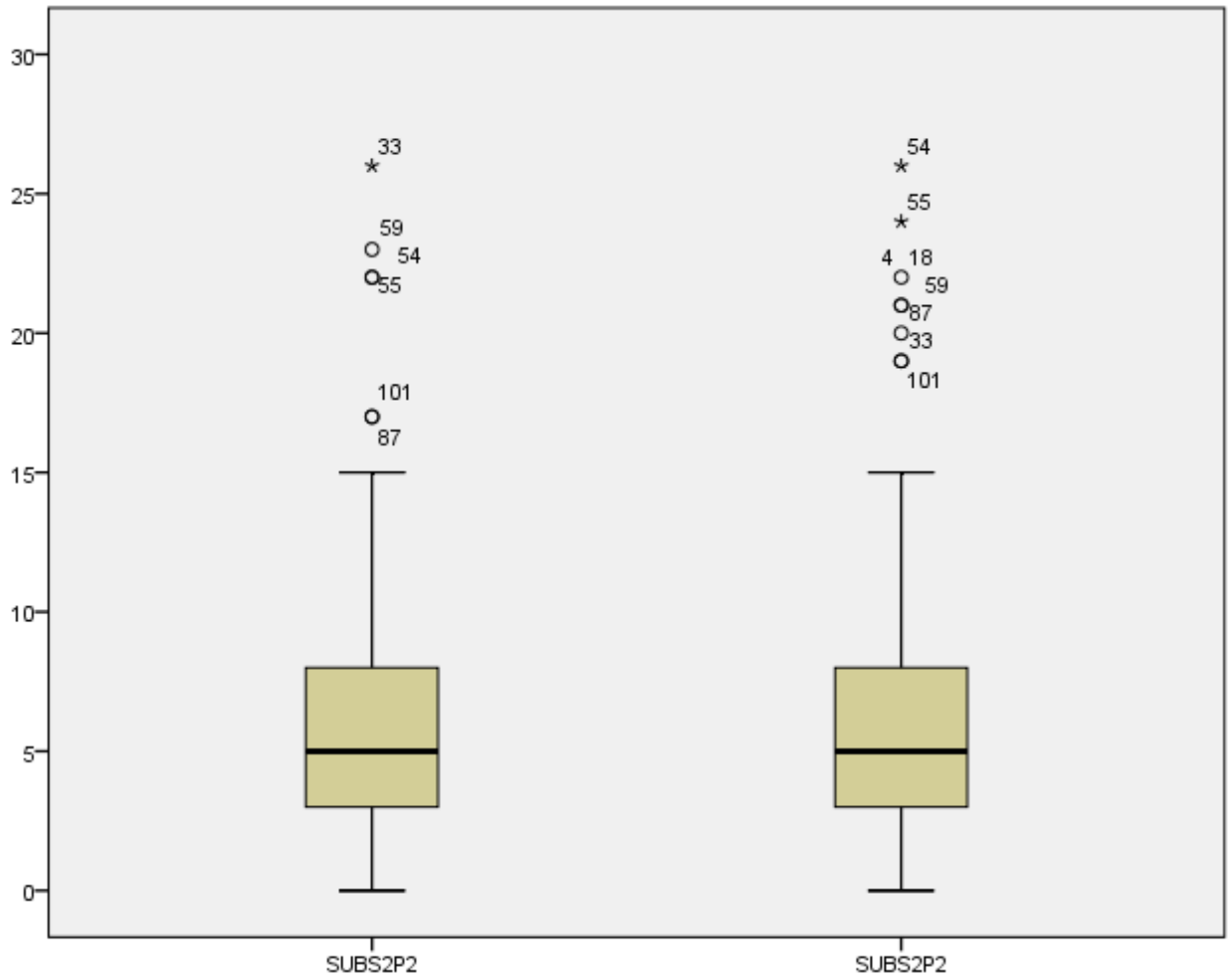
Figure 8: **Research Question 5 Section 2 Part 2 (S2P2) Outliers for error identification**

In the same fashion of the previous Research Questions 1, 3, and 4, the presence of outliers in the boxplots characterized the violation of normality assumption for research Question 5. Therefore, and in order to ensure the tests validity, as well as control for normality distribution, a series of tests were conducted to account for normality, skewness and kurtosis values. A review of the Shapiro-Wilk's test for normality in the pretest ($W=.837$, $p<.001$) showed statistical significance, which means that it did not meet the assumption for normality.

Both skewness (1.639) and kurtosis (2.786) statistics fell outside the acceptable range from -1.0 to +1.0, indicating violation of assumption of normality for the pretest.

Review of the Shapiro-Wilk's test for normality in the posttest (W=.848, p<.001) did not meet the assumption for normality. Both skewness (1.520) and kurtosis (2.168) statistics indicated violation of assumption of normality for the posttest, since these values were both outside the range of -1.0 to +1.0 for skewness and kurtosis. However, the -1.0 to +1.0 ranges are considered rather conservative, and some researchers would accept a wider range for skewness and kurtosis analyses. Despite these findings, and after removing the eight outliers present in the pretest, t-tests were rerun and the results continued to show no statistically significant difference in Section 2, Part 2 (S2P2) scores from pre- (M= 4.91, SD= 3.62) to posttest (M= 5.14, SD= 3.62), t(92)=.814, p.<.418. Research Question 5 was the only one of the five research questions whose result was not statistically significant after the treatment.

The boxplot for Research Question 5 (before the removal of outliers) is rather compelling in illustrating the almost identical range of the results, with both pre- and posttest starting at zero score and reaching very high results, where the concentration of outliers is. The interquartile range is also very similar, and the median remained unaltered after the treatment, five points out of 30 for maximum points. These observations seem to corroborate the inferential statistical findings, which suggest that indeed the treatment was not effective for explanation of grammatical errors of Research Question 5.

Although other parts of the ALAT showed a wide range of results, the boxplot for research Question 5 illustrates the heterogeneous nature of the data probably in the most extreme

way, with participants having results as low as zero and as high as almost maximum points, and keeping that same trend after the treatment.

Supplemental Results

The issue of *Attendance* was not a part of the primary research questions for the present study. Nevertheless, a decision to investigate it as a supplemental result was made in order to determine if there was a statistically significant difference in overall scores between pre- and posttest when participants received specifically either five or six sessions, as compared to analyzing all participants altogether regardless of their *Attendance*.

The first paired-samples *t*-tests considered all the participants who received the *Treatment* without discriminating the number of classes they attended. After specifying a number of at least five classes attended (out of the six possible sessions of the treatment), as a mandatory requirement for a participant to be included in the data collection, 11 participants were removed from the data, out of all the participants (N=101), totaling a new number of the participants (N=90) considered as full cases: pretest, treatment and posttest.

Paired samples *t*-tests were conducted after removing all the participants who did not attend at least five classes of the total of six classes of the full-treatment (*N*=90). For easy visualization of the results after the removal of the 11 participants, see Table 9. The analyses showed that the overall results for Research Question 1 remained statistically significant, showing increase in overall scores from pre to posttest.

Subsequent paired *t*-tests were also conducted for the other Research Questions 2, 3, 4, and 5. All of them were statistically significant with the exception of Research Question 5, not statistically significant before removing the 11 participants with poor attendance. Possibly

99

because the number of participants who did not take either five or six classes was so small, the overall results for both series of *t*-tests analyses were equally statistically significant, whether it included only participants at the ideal number of classes attended (at five or six), or when it included the 11 participants who had less than five classes.

Table 9: Results from pre- to posttest after the removal of 11 participants

Research Question	Maximum Points		Mean	Standard Deviation	N	*t*(89)	*p*
RQ 1 (Overall scores)	81	Pre	33.86	12.83	90	9.352	<.001
		Post	40.76	12.83			
RQ 2 (S1P1)	28	Pre	11.75	6.93	90	7.977	<.001
		Post	15.88	6.73			
RQ 3 (S1P2)	8	Pre	3.74	1.80	90	7.541	<.001
		Post	5.66	1.87			
RQ 4 (S2P1)	15	Pre	12.51	1.80	90	3.145	<.002
		Post	12.94	1.53			
RQ 5 (S2P2)	30	Pre	5.85	5.16	90	1.259	<.211
		Post	6.26	5.59			

Factorial ANOVAs were conducted to test whether or not *Attendance,* at either five or six classes attended*,* moderated the relationship between the *Treatment* and improvement in metalinguistic awareness (*N*=90). Consequently, there were two conditions for *Attendance* at either five or six classes attended. The ANOVA was not significant, $F(1, 88)=.1.072$, $p=.303$, which means that the change between pre- and posttest scores did not depend on participants attending five or six classes. In other words, if participants attended five or six classes, the *Treatment* effects were the same.

Instructor was not a part of the primary research questions for the present study, but a decision to investigate it as a supplemental result was made. Since there were two instructors providing the *Treatment*, it is possible that the effects of the treatment could vary depending on who was teaching the course content. Therefore, a factorial ANOVA was conducted to look at the interaction between *Instructor* and time of observation to determine if *Instructor* moderated the change between pre- and posttest on *Metalinguistic Awareness*. The interaction from the ANOVA was not significant $F(1, 88)=.404$, $p=.526$, which means that the change between pre- and posttest scores did not depend on which *Instructor* taught the classes. In other words, no matter which Instructor provided the *Treatment*, its effects were the same.

It is relevant to note that these findings point to what had been referred to previously as one of the controls for internal validity. In accordance with the notions of internal validity as proposed by Shadish et al. (2002), by ensuring that both instructors followed the same premises that the principal researcher had established, it was expected that the results would not show differences regarding which instructor taught the courses. The two instructors primarily: (a) taught the grammar same topics; (b) used similar presentations of content; (c) gave the same assignments (which in turn contained the same grammar topics). For this reason, the grammar module was very similar for both groups of participants who had the two different instructors. By the same premise, by ensuring similar treatment delivery, the study results hope to offer the possibility of accurate replication of the study and consequently generalizability of results across instructors.

CHAPTER 5:

CONCLUSION

The present empirical investigation confirmed previous research findings about preservice teacher preparation in relation to MA: preservice teachers are not receiving enough language-oriented exposure during their preparation years. Moreover, although improvement is feasible in a short period of time for more superficial aspects of language, the same improvement is not attainable in regards to more complex levels of language knowledge. Preservice teachers demonstrated a clear language -knowledge deficit while in college, and risk leaving it without significantly altering their language knowledge base and MA.

The issue of teacher preparedness has always been the focus of widespread discussions, and an essential topic among these discussions is the importance of language knowledge as a paramount factor that must be better accounted in preservice teachers' preparation (Myhill et al., 2013; Nutta et al., 2012).The understanding that preservice teachers need a more language-oriented approach during their preparation was the main factor that triggered the current study.

Research points to the need for preservice teachers to be better prepared in order to handle the numerous challenges of dealing with their future students. There seems to be a great concern and extensive literature endorsing more robust language knowledge on the part of the teachers. In their book *Preparing Every Teacher to Reach English Leaners,* Nutta et al. (2012) offered a deep look at both the statistics concerning the linguistically diverse population in the United States, as well as the solid regulations available to ensure that this population will receive adequate instruction.

According to Nutta et al., the number of ELs in the United States has reached the 5 million mark, and this number continues to grow. Federal legislation, including the No Child Left Behind Act of 2001 (NCLB), ensures that all institutions that offer education to ELs are rigorously supervised and held accountable for the instruction they receive, with clearly defined measures to guarantee such accountability. Two examples of how these institutions must comply with the NCLB are Title I and III provisions, which ensure that adequate yearly progress is made in reading and mathematics, and that annual achievements objectives are met. The Florida Department of Education determined that all the institutions that offer teacher education in the state of Florida must develop knowledge and skills specifically aimed at teaching English Language Learners (ELs). In order to fulfill such requirement, all preservice teachers who will teach language arts as part of their subject areas develop those skills during their teacher preparation courses.

However, and as confirmed by the findings in the present study, teacher preparation is still largely deficient in direct relation to how much language-oriented exposure preservice teachers are receiving. In the same vein, these findings point to the question of whether preservice teachers are having the opportunity to develop MA while they are still in college, during their prime preparation years.

The results of the current study demonstrated that *overall increase* of metalinguistic awareness is feasible, as attested by the results from Research Question 1. Nevertheless, further investigations (research Questions 2, 3, 4, and 5) gradually demonstrated that the level of metalinguistic awareness teachers vary significantly. These results echoed previous findings that demonstrated that preservice teachers are not proficient and language-knowledgeable enough to

deal with myriad issues that involve language, both in relation to students who are native speakers of English, but also concerning ELs (Kolln & Hancock, 2005; Nutta et al., 2012; Pappamihiell, 2007). Preservice teachers' lack higher level metalinguistic awareness, as evidenced by their limited ability to explain grammar issues and use proper metalanguage while doing so.

Research Implications

The present study points to a most basic premise: language is as pervasive as education itself, and as such it is the foundation of any teaching topic. More important, in order to be able to not only *use* language but also *understand* it from a more structural perspective, preservice teachers must be given the opportunity to develop and exercise this language understanding parallel to the development of their teaching in general.

Teachers must have robust language knowledge at *any level of teaching*, be it in the K-12, high school, adult education or English as a second language realm, to cite a few. The language knowledge that is being advocated through this investigation is one that empowers future teachers to control the narrative of *if*, *when*, and *how* they will need to explain grammar to their future students. It is only through an acquired expertise and confidence in how language works that preservice teachers will be able to gauge the right circumstances in which they use this language knowledge. Essentially, the notion being espoused is that grammar teaching should always be the result of a meaningful perspective (Celce-Murcia, 1992), and that this perspective must be based on clearly defined circumstance, and pedagogically-sound practices.

Preservice teachers should leave college knowledgeable enough to confidently provide more than superficial answers to their future students' language inquiries. They should have a

more structural understanding of language and embrace metalinguistic awareness before they enter the classroom. As pointed out by Folse (2016), teachers *should not be* in the classroom without this expertise, a view that is both espoused by the present study as well as demonstrated by its findings: institutions need to provide

Research question 1: Overall improvement is feasible. Research Question 1 asked whether there was a statistically significant increase between the overall scores of the ALAT, which measured the MA of the participants. The results from inferential statistical analyses showed a statistically significant increase from pre- to posttest. These results demonstrated that, even in a short period of time such as three weeks, it is possible to enhance metalinguistic awareness through explicit grammar instruction. Grammar topics are highly teachable and positive results can be reached in fairly short amounts of time, a notion that had been espoused by previous research findings (Cairns et al. 2004, 2006).

However, despite the statistical significance for most of the results (Research Questions 1, 2, 3, and 4), the part of the test that more specifically assessed the highest level of MA of participants (Research Question 5) was not statistically significant. This result demonstrated that ability to *explain* grammatical errors greatly differs from ability to *recognize* grammatical items or to *correct* grammatical errors. Also, descriptive results demonstrated great diversity in participants' performance percentages for each part of the test, pointing to caution while analyzing the results. Although more superficial notions of language can be easier to be achieved or improved, the same is not true for language issues of more complexity.

Although *t*-tests tend to be rather robust even with detection of outliers and consequently violation of assumption of normality, the existence of outliers for research Questions 1, 3, 4, and

5 required further investigation. However, more relevant to the discussion at hand, the aforementioned outliers were meaningful in order to better understand participants' range of knowledge. Participants' scores showed highly heterogeneous proficiency, with scores ranging from as low as zero to participants achieving either maximum or almost maximum scores. After further investigation of the background of the participants who were outliers in the Research Question 1(the overall scores), it is worth noting that, out of the four outliers with very high scores, two were participants who were either taking or had taken a specific grammar course in the recent past as part of their Teaching English as a Foreign Language (TEFL) certificate preparation. What this fact suggests is that any amount of extra exposure to grammar instruction can be perceived as a positive reinforcement to participants' previous knowledge, which in turn will lead to better performances. Furthermore, the aforementioned grammar course is solely dedicated to grammar, instead of having a great number of topics investigated more superficially.

Research question 2: Parts of Speech. Research Question 2 asked whether there was a statistically significant difference in the ability of participants to identify fourteen *parts of speech* after a treatment as measured by pre- and post-administrations of one of the parts of the ALAT (S1P1).

The results showed a statistically significant difference in recognition of parts of speech after the treatment, echoing similar findings in grammar performance after interventions in relatively short periods of time (Zipke, 2008, 2011). These results are even more relevant since it is known that a great number of these topics were covered in more detail during the treatment, reinforcing the notion that the treatment might have been directly responsible for the improved scores. The researcher's observations of the three-week grammar module (treatment) confirmed

106

that most of these items were explicitly taught during the treatment (see Table 2 on Grammar in Methodology).

A look at the descriptive statistics for some of the items offers a very diverse picture while observing the results for this part (see Figure 3). Items 1 and 5, (*verb and noun*) had above 80% of participants correctly identifying them in both pre- and posttests. On the other hand, Item 3 (*participle*) shows only 16% of participants scoring it correctly in the pretest and 12% in the posttest, a low result that finds echo in previous research studies for recognition of parts of speech (Bloor, 1986; Wray, 1993). In the same vein, Williamson and Hardman (1995) also pointed to relatively low results while testing 99 trainee primary school teachers: on a section assessing the ability to name parts of speech, participants achieved means scores of 5.6 out of 10.

As surprising (and perhaps disappointing) as it might be to find out that college students preparing to be teachers do not know primary parts of speech, it is important to reinforce the notion that the present study did find statistical significance from pre- to posttest for this research question, which means that here is reason for an optimistic view on grammar instruction as an efficient way of promoting knowledge in short amount of time. Descriptive statistics are merely offering an additional illustration for the inferential results, since they cannot provide support for claims of improvement or even difference at item level. These findings indicate that no matter the state of grammatical knowledge in which preservice teachers *initiate* their college preparation years, they can and *will improve* if given the opportunity to do so.

Research question 3: Parts of Sentence. Research Question 3 asked whether there was a statistically significant difference in the ability of participants to identify *parts of sentences*

after a treatment as measured by pre- and post-administrations of one of the parts of the ALAT (S1P2).

Results demonstrated that there was a statistically significant difference between pre- and posttests for research Question 3. A look at Figure 5 in the Results chapter offers an illustration of the four items that were tested for research Question 3, and its descriptive statistics. It is clear that *subject* had both the highest score in the pretest (91%), as well the highest score in the posttest (95%). However, this part of the test, with only four items, was the shortest and arguably easiest to score correctly even without being very confident of the right answers, since the ability to simply guess the results was obviously higher. After identifying the *subject*, which most participants seemed to be very proficient in doing, the chances of making a right choice between *direct* and *indirect object* were fairly easy. It is also relevant to reiterate that this section had the lowest Cronbach's alpha score of the entire test when analyzing parts individually ($\alpha.171$), which might reinforce the need to make it more balanced for future assessments. At least six options of sentences offering both *direct* and *indirect* object to be identified would make it more challenging to participants, and consequently provide more reliable results. Since the ALAT was an adaptation of a previously used test, the changes were minimal not to alter too many aspects of the instrument.

Research question 4: Correction of Grammatical Errors. Research Question 4 asked whether there was a statistically significant difference in the ability of participants to identify and correct common grammatical errors in a sentence after a treatment as measured by pre- and post-administrations of one of the parts of the ALAT (S2P1).

Results for this Research Question were not surprising, as research findings have previously pointed to correction of grammatical errors high accuracy, but failure to explain them (Alderson et al.1997). Figure 6 offers a descriptive glimpse of the very high scores (in the 80% to 100% range) for both pre- and posttest for most items of S2P1, with few exceptions (Items 3 and 13). Such results indicate that participants already knew very well how to correct these grammatical errors, a finding that echoes Alderson et al. (1997), whose research findings demonstrated that "students find it easy to correct errors in English" (p. 108). Nonetheless, this same ability is absent while trying to explain those very same errors.

In addition, although Research Questions 4 and 5 (identification and *correction* of errors and *explanation* of the same errors) are intrinsically related, and the sentences for both parts were exactly the same, their results were extremely different. Although the participants in this study (mostly native speakers of English), demonstrated that they knew how to detect and correct ungrammatical occurrences successfully, as demonstrated by the statistical significance of the analyses of Research Question 4, that ability did not translate well to explaining the same errors, which can be considered essential for them as future teachers (Folse, 2016).

Research question 5: Explanation of Grammatical Errors. Research Question 5 asked whether there was a statistically significant difference in the ability of participants to provide metalinguistic explanations for grammatical errors after a treatment as measured by pre- and post-administrations of one of the parts of the ALAT (S2P2). There was no statistical significance for the posttest after the treatment.

This research question encapsulates probably the most relevant aspect of the study. It confirms that mere ability to *correct* an ungrammatical sentence does not imply real

109

understanding of the error itself. The ALAT culminates with the most challenging activity within the test, and one that assesses MA at its highest level: ability to explain grammatical errors using adequate grammatical terminology. The results for this research question were dramatically different from the previous ones, pointing to a clear gap between ability to recognize terms and ability to explain grammatical issues. Participants consistently failed to demonstrate that they could explain grammatical errors, and do so using correct terminology. In addition to the inferential results, a look at the descriptive statistics (Figure 7), offers an additional illustration of participants' percentage scores starting as low as 7%, for *both* pre- and posttests, and not surpassing 43% in an overwhelmingly large number of questions. The treatment was clearly not efficient in promoting difference in participants' ability to go beyond mere recognition of terms to a level where they could name and explain types of grammatical errors.

The results for this particular part of the ALAT were not surprising. They confirmed the researcher's anecdotal experiences based on previous years teaching this same course. Although a great number of participants were able to accurately *correct* the grammar errors from the previous portion, they were clearly unable to correctly name and explain what type of error it was. This inability could be hypothesized as caused by (a) not enough instruction addressing more complex issues; and (b) a state of confusion on more complex aspects of grammar, to which participants were exposed but did not have time enough to process in the very short time between pre- and posttest.

Previous research findings consistently pointed to lack of proper metalanguage from teachers while analyzing language (Berry, 1997; Williamson and Hardman, 1995), and as stated before, such findings suggest a salient gap in the teacher preparation field. Preservice teachers

can neither adequately explain grammar topics nor can they correctly employ terminology about language while attempting to do so. Nonetheless, such ability to use language to depict, analyze , or explain language is an essential aspect of language proficiency for future teachers (Ellis, 2005; Myhill et al., 2013).

As stated by Folse (2016), this constitutes a serious issue for future teachers, who should not be in the classroom without this type of language expertise. In the same vein, while correction of errors relies on implicit knowledge, the ability to explain or teach these same errors would demand explicit knowledge of grammar rules and language as a system (Myhill et al., 2013), an ability that requires more than intuitive perception of correctness. Alderson et al. (1997) in a study that investigated applied linguistics students, offered the same view that they are not familiar with metalinguistic terms, despite being in a field that would supposedly predict such knowledge. Furthermore, when the college instruction offered in preparation years *assumes* an expertise that does not exist in reality, it will present serious challenges for students. In some cases, students' knowledge seems to be restricted to differentiate a *noun* from a *verb* (p. 108). Surely, most native speakers can intuitively correct an ungrammatical sentence in English. However, they cannot explain what the problem is with that sentence, and while this does not pose a threat to social interaction or functional use of language, it should be expected that future teachers are better prepared to offer these kinds of explanations.

Another interesting aspect of the results for Section 2, Part 2 (S2P2) , was that out of the eight high-end outliers present in the posttest, six were participants who had either taken or were taking a grammar course as part of their TEFL certification. Although inferential claims cannot be made based on this fact, it could be anecdotally interpreted as an indication that *the more*

grammar exposure students have, the better they will perform in an assessment of their grammar knowledge. A final note on parts four and five of the ALAT is that perspectives on what constitutes a grammatical error can be debatable and what is acceptable or not in terms of grammatical accuracy has changed in the last decades. Depending on one's views of prescriptive and descriptive grammar, some of these sentences are not necessarily wrong, or their ungrammatical element is not disruptive enough to be deemed relevant.

Results for Research Question 5 point to the fact that the amount of treatment that participants received was not enough to solidify more complex grammar knowledge such as explanation of errors and use of metalanguage. There is a compelling argument that there is a need for more language-focused training for preservice teachers. It is unacceptable that preservice teachers leave college unable to provide explanations for common grammar questions that any student in their future classrooms could one day ask.

Limitations

The first limitation of this study is the fact that the notions of MA encompass a lot more than the study was able to cover, both in in terms of the instrument used to measure this construct, as well as the treatment to which participants were exposed. Although the grammatical items in the instrument do represent a fair number of essential grammatical terms that any preservice teacher should be able to recognize and name, they certainly do not represent all the fundamental grammatical concepts which future teachers should know. The same can be said about the examples of ungrammatical sentences depicted in the open-ended sections of the test. Despite being fairly representative of what a future teacher should be able to recognize as

problematic, while also being able to explain them, these issues are not all there is to know about common grammar errors.

Another obvious limitation of the present study is the number of participants and the fact that its generalizability is limited to the immediate population from which the sample was extracted, namely the specific course within the institution where the study took place. A larger sample size, involving not only multiple semesters but also different institutions would offer a more compelling view of the current state of teacher preservice preparation today.

A final limitation of the study is that it offers only a quantitative picture of the results. There is a lot more to the possibilities that comprehend a period of three weeks in a college class than the numbers alone can provide. Issues such as motivation, aptitude, studying habits, and prior knowledge could bring a more detailed depiction of the results and offer some insight into what made the treatment successful or not to participants. Because of that, a combination of quantitative as well as qualitative data could offer a more thorough view of the current state of preservice teachers' language preparation, a notion espoused by multiple research findings (Harper & Rennie, 2008; Kolln &Hancock, 2005; Pappamihiell, 2007).

Future Research

Two recommendations for future research were drawn after the conclusion of this study: First, similar empirical studies should take place on a more continuous fashion, becoming a part of a standardized stage of any teacher preparation program. Second, a more comprehensive and balanced instrument should be designed specifically to evaluate preservice teachers' MA.

An assessment similar to the ALAT should become a routine diagnostic stage during preservice teachers' preparation, and as such it should be administered at least once during their

first two years of college. This will help increase their awareness of the *importance* of language and give them time to improve their language knowledge while they are still in college. By expanding the empirical knowledge of preservice teachers' MA and its ramifications in the teacher preparation field, educational institutions would be better informed on how to offer pedagogically-sound language courses for their teacher candidates

In regard to the instrument that was used for the present study, the Language Awareness Test (from which the present ALAT was primarily adapted) has been used for decades in various studies with multiple formats, and it is quite effective as an initial look at one's grammar knowledge and MA. However, it can be argued that it presents little balance among its own parts, at least in the format that was used for this study, and that it should be more updated to current views on grammar. For example, while some parts contain 14 or more items each, one of the parts has only. Cronbach's alpha analyses pointed to a low internal consistency for this part of the test; a more updated instrument will be a positive contribution to the teacher preparation field.

Conclusion

The current study's results demonstrate that even basic notions of grammar failed to be correctly explained, and that most participants were not able to use correct grammatical terminology to address the same issues. Moreover, although overall results showed improvement for recognition of terms, regarding a more sophisticated ability to name, explain and employ grammatical terminology, the study's findings demonstrated that a period of three weeks is not enough to ensure statistically significant difference from pre- to posttest.

The findings point to clear conclusions: (a) explicit grammar training can generate positive results after a treatment in a relatively short period of time; (b) improvement at *recognition* types of grammar exercise are better achieved than higher level *production* of grammar explanations after a treatment, in a relatively short period of time; (c) preservice teachers can accurately *correct* most ungrammatical sentences, however they lack both grammar terminology and proper metalanguage to *explain* grammatical errors.

A stronger preservice teacher language preparation. When grammar, syntax and morphology are receiving six classes *combined* in a course with an average of 32 classes, that speaks directly to the way the teacher preparation curriculum is approaching grammar: one topic among more than a dozen others. Furthermore, this is the only time in which grammar will be addressed for most preservice teachers during their preparation years. Except for those majoring in Language Arts or taking part in a TEFL certificate program, they will not have any other real opportunity to learn grammar. That is an obvious underestimation of the role of grammar and its importance to preservice teachers in their future careers.

The applied linguistics course in which participants were enrolled at the time of the study comprehends a great number of topics, ranging from the way language occurs in the brain to sociolinguistics. Each and every one of these topics is extremely relevant to their future careers as teachers, and it is undeniable that a course that gives preservice teachers an overall view of essential notions of language is not only welcome but necessary. Nevertheless, *grammar* is one among these topics, a worrisome realization which obviously restricts the *depth* in which it will be approached.

Preservice teachers' MA deserves a more prominent place in the current educational curriculum, and more should be done to emphasize it (Harper & Rennie, 2009; Myhill, 2000; Myhill et al., 2013). Instead of being fearful and insecure about how much (or how little) language knowledge they have, and how capable (or powerless) they are in order to handle their students' questions and difficulties about grammar, preservice teachers should be encouraged to discover the benefits of investigating their own language. More important, they should be able to do that while in college, during their preparation years, when they are starting to shape their own beliefs toward teaching. As pointed out by Celce-Murcia (1992), the notion of grammar teaching should not be detached from meaningfulness, but carried out in purposeful, productive ways. It is in this vein that the current study proposes an approach that invites preservice teachers to reflect upon their own MA and to work on improving their language skills in general.

Teachers' accountability toward their students' performances in general has never been so widely scrutinized, and such accountability has become a source of anxiety and concern for education professionals (Nutta et al., 2012; Pappamihiel, 2007). By acknowledging preservice teachers' own language limitations, institutions preparing them can improve the pathway for more effective practices in their years before entering a classroom. By setting a tone of confidence in how preservice teachers can improve their language knowledge while in preparation for their future careers as educators, this study proposes that instead of *assuming* that preservice teachers know enough about language, institutions should instead *assess* their knowledge and offer opportunities for improvement.

Preservice teachers are graduating with minimal metalinguistic knowledge (Harper &

Rennie, 2008), and this is a reality that must be changed. If little emphasis is given to the

language aspects of preservice teachers' preparation on the part of the institutions responsible for

them, such attitude might imply, even if involuntarily, an indifference to the one element that is

everywhere across the spectrum of education, teaching and learning: language. As proposed by

Kolln and Hancock (2005), instead of persevering on what seems to be (unacceptable) limited

amount of preparation regarding grammar, institutions preparing future teachers should be

convinced that lack of grammar is always more harmful than knowledge about it, and that the

more preservice teachers understand language, the better they will be able to help their future

students.

The understanding of how language is structured and the development of MA are

essential precursors for preservice teachers, and as such they should neither be ignored nor

underestimated.

APPENDIX A:

IRB APPROVAL LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Exempt Human Research

From: **UCF Institutional Review Board #1**
      FWA00000351, IRB00001138

To: **Aimee C. Schoonmaker**

Date: **January 23, 2015**

Dear Researcher:

On 01/23/2015, the IRB approved the following activity as human participant research that is exempt from regulation:

Type of Review:	Exempt Determination
Project Title:	INCREASING METALINGUISTIC AWARENESS AND LANGUAGE KNOWLEDGE AS NECESSARY PRECURSORS FOR PRESERVICE TEACHERS
Investigator:	Aimee C Schoonmaker
IRB Number:	SBE-15-10958
Funding Agency:	
Grant Title:	
Research ID:	N/A

This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these changes affect the exempt status of the human research, please contact the IRB. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 01/23/2015 04:33:40 PM EST

IRB Manager

APPENDIX B:
ALAT - ADAPTED LANGUAGE AWARENESS TEST
SECTION 1: GRAMMATICAL CATEGORIES

PART 1 - From the sentence below select **one** example of the grammatical item requested and underline write the item in the space provided in the two columns below.
Note: You may select the same word(s) **more than once** if appropriate:

> *Materials are delivered to the factory by a supplier, who usually has no technical knowledge but who happens to have the right contacts.*

1 - verb (any example)	8 - relative pronoun
2 - auxiliary verb	9 - adjective
3 - participle verb	10 - article: definite
4 - infinitive verb	11 - article: indefinite
5 - noun (any example)	12- adverb
6 - countable noun	13- preposition
7 - uncountable noun	14- conjunction

PART 2 - In the following sentences, underline the item requested in the second column:

1 - Poor little Joe stood out in the snow.	SUBJECT
2 - Joe has nowhere to shelter.	PREDICATE
3 - The policeman chased Joe down the street.	DIRECT OBJECT
4 - The woman gave him some money.	INDIRECT OBJECT

SECTION 2: GRAMMATICAL ERROR IDENTIFICATION

This section consists of fifteen English sentences, each of which contains a grammar mistake.
For each sentence, complete both **Parts 1** and **2**:
PART 1 – <u>Underline</u> the wrong part, and after the word **Correction**, write the sentence correctly.
PART 2 – After the word **Explanation**, <u>explain</u> the mistake.

1	I walk to work very quick. **Correction:** _____ **Explanation**: _____ _____
2	When her said that, Jack hit her. **Correction**: _____ **Explanation**: _____ _____
3	Every day I am making good resolutions. **Correction**: _____ **Explanation**: _____ _____
4	She's the taller of the four sisters. **Correction**: _____ **Explanation**: _____ _____
5	I live in a room at a top of an old house. **Correction**: _____ **Explanation**: _____ _____
6	Mommy goed to the park yesterday. **Correction**: _____ **Explanation**: _____ _____
7	The children put on their coat. **Correction**: _____ **Explanation**: _____ _____

8	He usually like to study at the library. **Correction**: _____ **Explanation**: _____ _____
9	I don't like people which are always apologizing. **Correction**: _____ **Explanation**: _____ _____
10	I opened the door, but I couldn't see nobody. **Correction**: _____ **Explanation**: _____ _____
11	When I was a small baby I have earaches. **Correction**: _____ **Explanation**: _____ _____
12	I will pick up you later. **Correction**: _____ **Explanation**: _____ _____
13	Josh and Pete have went to the show. **Correction**: _____ **Explanation**: _____ _____
14	Give the paper to Joe and I. **Correction**: _____ **Explanation**: _____ _____
15	She has called a few minutes ago. **Correction**: _____ **Explanation**: _____ _____

Thank you for taking this test!

APPENDIX C:

AUTHORIZATION LETTERS

Prof. S. Andrews
[sandrews@hku.hk]

Dear Aimee
 As promised, I am attaching the test and the mark scheme. Apologies for the delay.

When I have administered the test in the past, I have added a questionnaire to elicit test-takers' bio-data relevant to the independent variables I wanted to investigate in the study (e.g. to what extent does test performance vary according to years of teaching experience, or the relevance of the first degree? etc etc). You may wish to do the same.

Two important requests:

1.   Please make sure that you administer the test under exam conditions, and please don't allow anyone to take away/retain a copy of the test-paper etc.

2.   Also, in any report that you write – for your thesis or subsequently – please make sure that you acknowledge the test appropriately. You should say that you are using the test referred to in Andrews (2007). That test is largely based on Alderson, Clapham and Steel (1996), which in turn drew on Bloor (1986). The three references are given below.

Good luck! And if you have any further questions – about the test or the mark-scheme – just let me know.
 Best wishes
 Steve Andrews
 References:

Alderson, J.C., Clapham, C. and Steel, D. (1996). *Metalinguistic knowledge, language aptitude and language proficiency*, Working Papers, 26, University of Lancaster.

Andrews, S.J. (2007). *Teacher language awareness*. Cambridge: Cambridge University Press.

Bloor, T. (1986). What do language students know about grammar? *British Journal of Language Teaching*, 24(3), 157-60.

Prof. C. Alderson
[c.alderson@lancaster.ac.uk]

By all means adapt the instrument to your needs, and you simply acknowledge the original source when you write about it.

The only permission you might need, but I don't remember having asked for permission myself, is the third section – Words in sentences – which came from the Modern Language Aptitude Test by John Carroll. I have now found a copy of that section and am attaching it.

By the way, have you seen the article I wrote with Richard Hudson. In case you haven't, I am attaching a copy.

Good luck with your study and be sure to send me a copy once you have published on it

Best wishes
Charles

APPENDIX D:

ALAT MARK SCHEM

SECTION 1: PARTS 1 & 2

Point Distribution: **Section 1 – Parts 1 & 2**	
Correct answer: 2	**Incorrect answer/no answer: 0**

Section 1 – Part 1			
Item	Possible answers	Item	Possible answers
1 - verb	*are, delivered, has, happens, have, to have*	8 - relative pronoun	*who*
2 - auxiliary verb	*are* (don't accept both *are delivered*)	9 - adjective	*technical, right*
3 - participle verb	*delivered* (don't accept both *are delivered*)	10 - definite article	*the* (don't accept both *the/a*) (don't accept both *a/the*)
4 - infinitive verb	*to have*	11 - indefinite article	*a* (don't accept both *the/a*) (don't accept both *a/the*)
5 - noun	*materials, factory, supplier, knowledge, contacts*	12 - adverb	*usually, no*
6 - countable noun	*materials, factory supplier, contacts*	13 - preposition	*to, by*
7 - uncountable noun	*knowledge*	14 - conjunction	*but*
		Total Points:	**28**

Section 1 – Part 2		
Item		Possible answers
1 - Poor little <u>Joe</u> stood out in the snow. <u>Poor little Joe</u> stood out in the snow.	Subject	*Joe, Poor little Joe*
2 - Joe <u>has</u> nowhere to shelter. Joe <u>has nowhere to shelter</u>.	Predicate	*has, has nowhere to shelter*
3 - The policeman chased <u>Joe</u> down the street.	Direct Object	*Joe*
4 - The woman gave <u>him</u> some money.	Indirect Object	*him*
	Total Points:	**8**

SECTION 2: PARTS 1 & 2

Point Distribution – **Section 2 – Part 1 – Correction**	
Correct answer: 1	**Incorrect answer/no answer: 0**
Point Distribution – **Section 2 – Part 2 – Explanation**	

Fully correct explanation: 2
<u>Rule with grammar terminology</u>: correct explanation that shows knowledge of the rule being broken using clear grammar terminology.
Partially correct explanation: 1
<u>Rule only or informal explanation of the rule</u>: correct, but informal explanation with minimal or no grammar terminology. It goes beyond simply *describing* the error but it doesn't use as much grammar terminology as the fully correct version.
Incorrect explanation: 0
Either no explanation or incorrect explanation. It simply describes the correction from Section 2 – Part 1.

Section 2 – Parts 1 & 2 - Explanations & Sample Corrections	

1	**Sentence:** I walk to work very quick.
	Correction: *I walk to work very **quickly**.*
	Fully correct explanation: 2 points
	Explains the rule or difference between adverbs and adjectives when describing or qualifying verbs. Do not accept removal of *very* as part of the error.
	Uses one or more of the following or similar terms: *verb, adverb, adverb of manner, adjective, qualifying/describing verbs, -ly, morpheme, suffix.*
	e.g., When a verb is qualified, use adverb (not adjective).
	    Quick is an adjective; it must be an adverb to qualify a verb.
	    Add the morpheme *-ly* to the adjective so it becomes an adverb, and it refers to verb *walk*.
	
	Partially correct explanation: 1 point
	Some awareness of rule and/or metalanguage, but no explanation.
	e.g., *-ly* for adverb.
	    Missing adverbial ending.
	    Quick is an adjective, *quickly* is an adverb.
	
	Incorrect explanation: 0
	Simply describes the correction from previous Part 1. It does not explaining anything, just restates the obvious.
	e.g., *Quick* should be *quickly*.
	    Wrong adverb.
	    Adjective issue.

2	**Sentence:** When her said that, Jack hit her. **Correction:** *When **she** said that, Jack hit her.* *When **he** said that, Jack hit her.* *Jack hit her when **she** said that.*

	Fully correct explanation: 2 points
	Explains the rule or difference between object and subject pronouns.
	Uses one or more of the following or similar terms: *subject, object, pronouns, personal pronouns, subject pronoun, object pronoun, possessive adjective, subjective case, objective case.*
	e.g., The pronoun must be used in the subjective, not objective case.
	    The pronoun *her* cannot be used as the subject of the sentence.
	   Her can be either an object pronoun or a possessive adjective; it cannot be in the position of subject pronoun performing actions.
	
	Partially correct explanation: 1 point
	Some awareness of rule and/or metalanguage, but no explanation.
	e.g., The word before the verb does the action, *she*, not *her*.
	
	Incorrect explanation: 0
	Simply describes the correction from previous Part 1. It does not explaining anything, just restates the obvious.
	   e.g., *Her* should be *she*.
	       Her is not an appropriate word.
	       Wrong pronoun.

3	**Sentence:** Every day I am making good resolutions **Correction:** *Every day I **make** good resolutions.*

	Fully correct explanation: 2 points
	Explains the rule or difference between simple present tense for habitual actions and/or present continuous/progressive for an action happening at the moment of speaking.
	Uses one or more of the following or similar terms: *tense, present tense, progressive/continuous tense, habitual actions in the present, prescriptive/descriptive grammar*.
	e.g., Simple present tense should be used for habitual/everyday actions and events.
	     Present continuous or progressive tense is not adequate for habitual actions in the present.
	
	Partially correct explanation: 1 point
	Shows some awareness of rule and/or metalanguage, but no explanation.
	e.g., Should be simple present tense.
	   We are speaking of *every day*, not *right now*.
	
	Incorrect explanation: 0
	Simply describes the correction from previous Part 1. It does not explaining anything, just restates the

	obvious. e.g., Tense issue. 　　Wrong tense.
4	**Sentence**: She's the taller of the four sisters. **Correction**: *She's the **tallest** of the four sisters.*
	Fully correct explanation: 2 points Explains the rule or difference between comparing two things/people or more than two. Uses one or more of the following or similar terms: *comparative adjectives, superlative adjectives, -ed and -est morphemes, suffixes -er, -est, morphology.* e.g., When comparing more than two things/persons, use superlative adjective, not comparative. 　　*Taller* is a comparative adjective, *tallest* is the correct superlative adjective to compare more than two things/people. 　　*Taller* is a comparative used to compare two or more things, *tallest* is a superlative used to set one thing above more than two. 　　The right morpheme/suffix/ particle to be added to the adjective *tall* is *–est*, not *–er*. **Partially correct explanation: 1 point** Shows some awareness of rule and/or metalanguage, but no explanation. e.g., Two things, use *–er*, more than two things, use *-est*. 　　Comparative to superlative. 　　A superlative is needed. **Incorrect explanation: 0** Simply describes the correction from previous Part 1. It does not explaining anything, just restates the obvious. e.g., Wrong suffix. 　　*Taller* to *tallest*. 　　Wrong ending.
5	**Sentence:**　I live in a room at a top of an old house **Correction:** *I live in a room at **the** top of an old house.*
	Fully correct explanation: 2 points Explains the rule or difference between definite and indefinite articles*,* or why indefinite article is incorrect in this case. Uses one or more of the following or similar terms: e.g., Use definite article *the* to refer to something specific. 　　Definite article *the* is used when referring to only one, specific item; article *a* is indefinite. **Partially correct explanation: 1 point** Shows some awareness of rule and/or metalanguage, but no explanation. e.g., It needs to be a definite article. 　　There is only one top of the house; it's a specific place.

	Incorrect explanation: 0 Simply describes the correction from previous Part 1. It does not explaining anything, just restates the obvious. e.g., Wrong article. *A* should be *the*.
6	**Sentence:** Mommy goed to the park yesterday. **Correction**: *Mommy **went** to the park yesterday.*

	Fully correct explanation: 2 points
	  Explains why *goed* is not correct because of the irregularity of the verb *to go* in the past.
	  Uses one or more of the following or similar terms: *past tense, irregular/regular verbs in the past, overgeneralization of the –ed rule, verb form, -ed morpheme, -ed suffix, morphology.*
	e.g., The verb *to go* is irregular. Therefore, its past form is *went*, not *goed*.
	    Incorrect verb form, the verb *go* is an irregular verb.
	    The past of verb *to go* is not formed by adding –*ed* to its end, like other regular verbs.
	    Most verbs form their past by adding –*ed*, but not *go*; overgeneralization of rule.
	
	Partially correct explanation: 1 point
	Shows some awareness of rule and/or metalanguage, but no explanation.
	e.g., *Goed* is not a word, the past tense of *go* is *went*.
	    This is an *overextension* issue.
	        • If they wrote only that, MARK 0.
	        • If they explain it correctly, MARK 1. They mean *overgeneralization*, but these two terms are easily confused, so if the rationale is correct it can be accepted.
	
	Incorrect explanation: 0
	Simply describes the correction from previous Part 1. It does not explaining anything, just restates the obvious.
	e.g., Wrong verb ending.
	    Goed should be *went*.
	    Goed is not a word (without explaining).
	    Past.

7	**Sentence:** The children put on their coat. **Correction:** *The children put on their **coats**.*
	Fully correct explanation: 2 points Explains the specific relationship in this sentence of the subject *children* and the direct object *coats*. In this case there is a need for agreement between subject and direct object because the children are wearing more than one coat. Uses one or more of the following or similar terms: *subject, object, number agreement, subject/object agreement, plural of noun, plural of children.*

e.g., The subject *children* is plural. Therefore, the direct object must agree with the subject.

The plural subject *children* requires the object *coat* to agree in number, becoming *coats*.

Children is a plural noun, and they are wearing the coats. Therefore, the direct object *coats* must be also plural

Each child is wearing one individual coat. Therefore, the noun *coat* must agree with the subject *children*.

Partially correct explanation: 1 point
Shows some awareness of rule and/or metalanguage, but no explanation.
e.g., The children are not sharing the same coat, so coat must be in plural.

If there is more than one child there must be more than one coat.

Incorrect explanation: 0
Simply describes the correction from previous Part 1. It does not explaining anything, just restates the obvious.
e.g., Wrong ending for coat.

Plural matching.

More than one child.

More than one coat.

8	**Sentence:** He usually like to study at the library. **Correction:** *He usually **likes** to study at the library;*

Fully correct explanation: 2 points
Explains the subject-verb agreement rule: verb needs to agree with the subject. In this case, the third-person singular and the addition of –*s* at the end of the verb.
Uses one or more of the following or similar terms: subject-verb agreement, subject and verb, simple present tense, third-person singular, add -*s* to the verb.
e.g., Present tense of the verb *like* needs to agree with the subject *he*, which is a third-person singular.

He is one of the third-person singular pronouns and it requires addition of -*s* to the end of the verb in the simple present tense.

Subject-verb agreement issue: verb needs to agree with the subject.

Partially correct explanation: 1 point
Shows some awareness of rule and/or metalanguage, but no explanation.
e.g., Subject-verb agreement.

3rd person singular rule

Incorrect explanation: 0
Simply describes the correction from previous Part 1. It does not explaining anything, just restates the obvious. This is not a tense issue, since *like* is also present. This is not a number issue either; there is no singular/plural of the verb but of the pronoun. It is not explaining anything, just restating the obvious
e.g., Tense issue.

Like is plural.

Pluralization of verb.

9	**Sentence**: I don't like people which are always apologizing. **Correction**: *I don't like people **who** are always apologizing.* 　　　　　　 *I don't like people **that** are always apologizing.*
	Fully correct explanation: 2 points 　Explains the difference between people and things when using a relative pronoun in a restrictive relative clause. 　Uses one or more of the following or similar terms: relative pronoun, pronouns, restrictive clauses. e.g., The relative pronoun *who* should be used instead of *which* to refer to people in a restrictive clause. 　　Use pronouns *who/that* for people and *which* for things. **Partially correct explanation: 1 point** Shows some awareness of rule and/or metalanguage, but no explanation. e.g., *Who* for people. 　　*Which* for things. 　　*Who/that* for people, *which* for things. **Incorrect explanation: 0** Simply describes the correction from previous Part 1. It does not explaining anything, just restates the obvious. e.g., *Who*, not *which*. 　　Has to be *who*. 　　*Which* is wrong.
10	**Sentence**: 　I opened the door, but I couldn't see nobody. **Correction**: *I opened the door, but I couldn't see **anybody** or **anyone*** 　　　　　　*I opened the door, but I **could** see nobody.* **Fully correct explanation: 2 points** 　Explains the double-negative issue as ungrammatical in standard English. Uses one or more of the following or similar terms: *double-negative, non-standard English, ungrammatical, descriptive, prescriptive grammar.* e.g., A double-negative sentence is wrong if the meaning is intended as negative. 　　*Couldn't* and *nobody* make it a double negative. 　　Double-negative sentences become positive; it's wrong in English but correct in other languages. **Partially correct explanation: 1 point** Shows some awareness of rule and/or metalanguage, but no explanation. e.g., If you can't see anybody, you can see somebody. 　　Prescriptive grammar issue. 　　Double negative issue. **Incorrect explanation: 0**

	Simply describes the correction from previous Part 1. It does not explain anything, just restates the obvious.
	 e.g., Not *nobody, anybody*.
	     Nobody is not proper.
	     Wrong.

11	**Sentence:** When I was a small baby, I have earaches. **Correction**: *When I was a small baby, I **had** earaches;* *When I was a small baby, I **used to have** earaches.*

	Fully correct explanation: 2 points
	 Explains the rule for tense-agreement, or the need for both actions to be in the same tense; in this case, in the simple past tense.
	    Uses one or more of the following or similar terms: *tense agreement, simple past, completed actions, definite time in the past, main clause. If the correction included 'used to have', an expected term is 'habitual action'.*
	 e.g., The two sentences/ clauses must agree in the simple past tense.
	     The dependent clause must agree in tense with the main clause.
	     When I was indicates past tense, *I have earaches* is present, there is no agreement between the two.
	
	 Partially correct explanation: 1 point
	 Shows some awareness of rule and/or metalanguage, but no explanation.
	 e.g., Past tense must agree with previous sentence.
	     I was means past, so *I have earaches* must be past.
	     You are not a small baby anymore, so you can't say *I have earaches* in the present.
	
	 Incorrect explanation: 0
	 Simply describes the correction from previous Part 1. It does not explain anything, just restates the obvious.
	 e.g., Wrong tense.
	     Change tense.
	     Past.

12	**Sentence:** I will pick up you later. **Correction:** *I will **pick** you **up** later.*
	Fully correct explanation: 2 points Explains that this is an example of a transitive, separable phrasal verb. It must be separated to include the direct object. Uses one or more of the following or similar terms: *phrasal verb, separable, transitive, direct object, object pronoun, preposition, particle.* e.g., The phrasal verb *pick up* is transitive and separable; the object must be positioned between the verb *pick* and the preposition *up*. The phrasal verb *pick up* must be separated to include the direct object/object pronoun. **Partially correct explanation: 1 point**

Shows some awareness of rule and/or metalanguage, but no explanation.

e.g., Word order of phrasal verb.

Pick and *up* must split to accommodate the pronoun *you*.

The direct object *you* follows the verb.

Incorrect explanation: 0

Simply describes the correction from previous Part 1. It does not explain anything, just restates the obvious.

e.g., Switch *pick* and *up*.

Wrong order.

Wrong SVO.

13	**Sentence:** Josh and Pete have went to the show. **Correction:** *Josh and Pete have **gone** to the show.* *Josh and Pete **went** to the show (Partially correct: accept only if the explanation is valid.*

	Fully correct explanation: 2 points
	Explains the formation of present perfect tense, and the lack of past participle for the main verb.
	Uses one or more of the following or similar terms: *present perfect tense, past participle, past tense, main verb, irregular main verb, auxiliary verb.*
	e.g., Present perfect tense formation requires main verb to be in the past participle, not past form.
	    When using present perfect tense, the main verb must be in the past participle form.
	    The main verb *go* must be in the past participle form, not past.
	
	Partially correct explanation: 1 point
	Shows some awareness of rule and/or metalanguage, but no explanation.
	e.g., Present perfect tense formation.
	    The auxiliary *have* is not needed if relating to the past tense. (*Josh and Pete went to the show*)
	    Went already indicates simple past. (*Josh and Pete went to the show*)
	
	Incorrect explanation: 0
	Simply describes the correction from previous Part 1. It does not explain anything, just restates the obvious.
	e.g., Tense issue.
	    Have is present, *went* is past.
	    Improper formation.

14	Give the paper to Joe and I. **Correction:** *Give the paper to Joe and **me** or **me** and Joe.* *Give the paper to **us**.* *Give **us** the paper.*
	Fully correct explanation: 2 points Explains the rule for subject and/or object pronouns. Uses one or more of the following or similar terms: *subject pronoun, object pronoun, pronoun, objective case, subjective case, preposition.* e.g., *I* is a subject pronoun; it cannot be used in the position of (indirect) object.

It's wrong to use *I*, a subject pronoun, in the (indirect) object position; it needs to be replaced with the object pronoun *me*.

   Joe and I should be replaced with an object pronoun, not a subject pronoun.

   After prepositions, use objective case pronouns.

Partially correct explanation: 1 point
Shows some awareness of rule and/or metalanguage, but no explanation.
e.g., You wouldn't say: *Give the paper to I*, so you can't say: *Give the paper to Joe and I.*
   If you remove Joe it becomes *Give the paper to I*, that is how you know it's not correct.

Incorrect explanation: 0
Simply describes the correction from previous Part 1. It does not explain anything, just restates the obvious.
e.g., Wrong pronoun.
   Change *I* to *us*.
   Improper grammar.

15	**Sentence:** She has called a few minutes ago. **Correction:** *She **called** a few minutes ago.* *She **has called**.* (Without *a few minutes ago:* Only accept if the explanation is valid)

Fully correct explanation: 2 points
Explains the rule for complete action in the simple past tense or the difference between present perfect tense and simple past tense when using a time expression like *ago*, which indicates definite time in the past.
Uses one or more of the following or similar terms: *present perfect tense, past tense, past time expression, definite time in the past, indefinite time in the past, auxiliary verb has.*
e.g., When describing a complete past action, use simple past tense, not present perfect tense.
   Ago indicates a completed past action. Therefore, present perfect tense is wrong.
   Auxiliary *has* is not necessary if the tense is simple past.
   The action is in the past, but the time is not definite, so you can't use *ago*. (For the correction *She has called*)

Partially correct explanation: 1 point
Shows some awareness of rule and/or metalanguage, but no explanation.
e.g., Past, not present perfect tense.
   Action already happened, past.
   The sentence shows that the call already happened.

Incorrect explanation: 0
Simply describes the correction from previous Part 1. It does not explain anything, just restates the obvious.
e.g., *Has* is not necessary.
   Just remove *has*.
   It sounds correct.

TOTAL POINTS – SECTION 2 - PART 1	**15**
SECTION 2 - PART 2	**30**
TOTAL POINTS FOR SECTION 1 (PARTS 1 & 2) & SECTION 2 (PARTS 1 & 2)	**81**

APPENDIX E: DATA COLLECTION SCRIPT

Anyone proctoring the Pre-test will say:

- *The language activity you are about to do is very simple.*
  - *The Principal Researcher explained the study to you on February 24th, 2015, and you also received an Explanation of Research through your Canvas webcourses.*
  - *All the information you need to know about the study is in the Explanation of Research.*
  - *If you haven't read the Explanation of Research, it is available here for you.*

Show the sheet Explanation of Research.

- *REGARDLESS OF YOUR PERFORMANCE in the tests, you will receive 5 points for TSL4240 if you fill the bio-data questionnaire and take both pre- and posttests in this study.*
- *You can use pencil or pen to write your answers.*
- *Everyone in the room today will be doing this activity at the same time, so be careful not to get distracted by others, and do not talk to anybody while taking the test.*
- *You will have exactly 30 minutes to take the entire test, and no more than that.*
- *Please, fill the bio-data questionnaire at the beginning of the test.*
- *The activity will start AFTER EVERYONE has received the test.*
- *Please, turn off your phones.*
- *Thank you for participating in this study!*

Before handing the tests, ask participants to keep their tests face down on their desks until all participants have received their tests.

Start calling the names on the tests and handing the tests to participants. If there are two people proctoring the test, split the pile of tests in two.

Write down the time of beginning on the package and set the alarm for 30 minutes after that.
Say: *You can begin the test now.*
After 30 minutes, say: *Time is over, please submit your tests.*

REFERENCES

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language Test Construction and Evaluation.* Cambridge University Press, UK.

Alderson, J. C., Steel, D., & Clapham, C. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research, 1*(2), 93–121.

Alderson. J. C. & Hudson, R. (2012). The metalinguistic knowledge of undergraduate students of English language or linguistics, *Language Awareness,* 1-18.

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

Andrews, S. (1997). Metalinguistic awareness and teacher explanation. *Language Awareness*, *6*(2-3), 147-161.

Andrews, S. (1999a). Why do L2 teachers need to 'know about language'? Teacher metalinguistic awareness and input for learning. *Language and Education, 13*(3), 161-177.

Andrews, S. (1999b). 'All these like little name things': A comparative study of language teachers' explicit knowledge of grammar and grammatical terminology. *Language Awareness, 8*(3), 143-159.

Andrews, S. (2001). The language awareness of the L2 teacher: Its impact upon pedagogical practice. *Language Awareness*, *10*(2-3), 75-90.

Andrews, S. (2003). Teacher language awareness and the professional knowledge base of the L2 teacher. *Language Awareness*, *12*(2), 81-95.

Andrews, S. (2006). The evolution of the teachers' language awareness. *Language Awareness, 15*(1), 1-19.

Andrews, S.J. (2007). *Teacher language awareness*. Cambridge: Cambridge University Press.

Apel, K., Fowler, E., Brimo, D., & Perrin, N. (2012). Metalinguistic contributions to reading and spelling in second and third grade students. *Reading and Writing, 25*, 1283–1305.

Berry, R. (1997). Teachers' awareness of learners' knowledge: The case of metalinguistic terminology. *Language Awareness, 6*(2), 136-146.

Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition, 122*, 67-73.

Bialystok, E., & Ryan, E. (1985). Toward a definition of metalinguistic skill. *Merrill-Palmer Quarterly, 31*(3), 229-251.

Bloor, T. (1986). What do language teachers know about grammar? *British Journal of Language Teaching*, *24*(3), 157-162.

Borg, S. (2003).Teacher Cognition in Grammar Teaching: A Literature Review. *Language Awareness, 12*(2), 96-108.

Cairns, H. S., Waltzman, D., & Schlisselberg, G., (2004). Detecting the ambiguity of sentences: Relationship to early reading skill. *Communication Disorders Quarterly, 25*(2), 68-79.

Cairns, H. S., Schlisselberg, G., Waltzman, D., & McDaniel, D. (2006). Development of a metalinguistic skill. *Communication Disorders Quarterly, 27*(4), 213-220.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Boston, MA: Houghton Mifflin.

Celce-Murcia, M. (1992). Under what circumstances, if any, should formal grammar instruction take place? *TESOL Quarterly, 26*(2), 406-409.

Chomsky, N. (1968). *Language and Mind*. Harcourt, Brace and World, New York.

Clapham, C. (2001).The assessment of metalinguistic knowledge. In Elder, C., Brown, A. Grove, E., Hill, K., Iwashita, N. & Lumley, T. *Experimenting with uncertainty: Essays in honor of Alan Davies* (p.33-43). Cambridge: UCLES.

Crystal, D. (2006). (2006). *The Fight for English: How language pundits ate, shot, and left*. Oxford: Oxford University Press.

Davies, A., Brown, C., Hill, K., Lumley, T., McNamara, T. (1999). *Dictionary of Language Testing*. University Press, Cambridge, UK.

Donmall, B. G. (1985). Language awareness: NCLE reports and papers 6. London: Centre for Information on Language Teaching and Research.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning, 54*(2), 227-275.

Ellis, R. (2005). Measuring implicit and explicit knowledge of second language: A Psychometric Study. *SSLA*, 27, 141-172.

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA Perspective. *TESOL Quarterly, 40*(1)83-107.

Ellis, R. (2008). *The study of second language acquisition*. Oxford University Press.

Fielding-Barnsley, R., & Purdie, N. (2005). Teachers' attitude to and knowledge of

    metalinguistics in the process of learning to read. *Asia-Pacific Journal of Teacher*

    Education, *33*(1), 65-76.

Fielding-Barnsley, R. (2010). Australian pre-service teachers' knowledge of phonemic

    awareness and phonics in the process of learning to read. *Australian Journal of Learning*

    Difficultities, 15(1), 99-110.

Folse, K. (2009). *Keys to teaching grammar to English language learners*: *A practical*

    handbook. Ann Arbor: University of Michigan Press.

Folse, K. (2016). *Keys to teaching grammar to English language learners: A practical*

    handbook. (2nd ed.). Ann Arbor: Michigan University Press.

Galguera, T. (2011). Participant structures as professional learning tasks and the development of

    pedagogical language knowledge among preservice teachers. *Teacher Education*

    Quarterly, 3, 86-106.

Gamaroff, R. (2000). Rater reliability in language assessment: the bug of all bears. *System*, *28*,

    31–53.

Gregg, K. R. (1989). Second language acquisition theory: The case for a generative perspective.

    In S. M. Gass & J. Schachter (Eds.), Linguistics Perspective on second language

    acquisition (p. 15-40). New York: Cambridge University Press.

Green, J. B., & Salkind, N. J. (2011). *Using SPSS for Windows and Macintosh: Analyzing and*

    understanding data (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Harper, H., & Rennie, J. (2009). 'I had to go out and get myself a book on grammar': A study of pre-service teachers' knowledge about language. *Australian Journal of Language & Literacy*, *32*(1), 22-37.

Henshaw, A. (1985). Latin Rules! *British Journal of Language teaching, 23*(2), 69-71. ERIC EJ328213.

Hun, Y., Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research, 2*(1), 1-23.

Karmiloff, S. (1979). Micro- and macro-developmental changes in language acquisition and other representation systems. *Cognitive Science, 3*, 91-118.

Kolln, M. & Hancock, C. (2005). The story of English grammar in the US schools. *English Teaching: Practice and Critique* 4(3) 11-31.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford, England: Pergamon.

Krashen, S. (1982). *Principles and practice in second language acquisition*. New York: Prentice Hall.

Krashen, S. (1992). Under what circumstances, if any, should formal grammar instruction take place? *TESOL Quarterly, 26*(2), 409-411.

Krashen, S (1993). The effect of grammar teaching: Still peripheral. *TESOL Quarterly, 27*, 717-725.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational Psychology Measure, 30*(3), 607-610.

Lewkowicz, N., (1980). Phonemic awareness training: What to teach and how to teach it. *Journal of Educational Psychology, 72*(5), 686-700.

Lomax, R. G. (2007). *An introduction to statistical concepts for education and behavioral sciences (2nd ed.)*. Mahwah, NJ: Lawrence Erlbaum.

Moats, L. (1994). The missing foundation in teacher education: Knowledge of the structure of spoken and written language. *Annals of Dyslexia, 44*, 81-102.

Moats, L. (2009). Knowledge foundations for teaching reading and spelling. *Reading and Writing, 4(*22), 379–399. DOI 10.1007/s11145-009-9162-1

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*. Amsterdam: John Benjamins.

Myhill, D. (2000). Misconceptions and difficulties in the acquisition of metalinguistic knowledge. *Language and Education, 14*(3), 151-163. DOI: 10.1080/09500780008666787

Myhill, D. A., Jones, S. M. (2011). Policing grammar: The place of grammar in literacy policy. In A. Goodwyn, & C. Fuller. *The Literacy Game*, 44-62. London, Routledge.

Myhill, D. A., Jones, S. M., Lines, H., & Watson, A. (2012). Re-thinking grammar: The impact of embedded grammar teaching on students' writing and students' metalinguistic understanding. *Research Papers in Education, 27*(2), 1-28.

Myhill, D., Jones, S.M., & Watson, A. (2013). Grammar matters: How teachers' grammatical knowledge impacts on the teaching of writing. *Teaching and teacher Education, 36*, 77-91.

NCLE (1985). National Council on Languages in Education: Ed. Donmall, Gill. *Language
Awareness*: NCLE Papers and Reports 6, CILT.

No Child Left Behind (NCLB) Act of 2001, 20 U.S.C.A. § 6301 *et seq.* (West 2003).

Norton, B., & Toohey, K. (2004). *Critical Pedagogies and language learning*. New York:
Cambridge University Press.

Nutta, J., Mokhtari, K., & Strebel, C. (2012). *Preparing every teacher to reach English learners:
A practical guide for teacher educators*. Cambridge, MA: Harvard.

O'Grady, W, Archibald, J. Aronoff, M. & Rees-Miller, J. (2010). *Contemporary Linguistics: An
Introduction*. Bedford St. Martin's. Boston.

Pappamihiel, E. (2007). Helping preservice content-area teachers relate to English language
learners: An investigation of attitudes and beliefs. *TESL Canada Journal, 24*(2), 42-60.

Roehr, K., & Ganem-Gutierrez, G. (2009). The status of metalinguistic knowledge in instructed
adult L2 learning. *Language Awareness, 18*(2), p. 165-181.

Salkind, N. J. (2008). *Statistics for people who (think they) hate statistics (3rd ed.).* Los Angeles:
Sage Publications.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental
designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Shavelson, R. J., & Towne, L. (2002). *Scientific research in education*. Washington, DC:
National Academies Press.

Skinner, B.F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

TESOL, *ESL Standards for Pre-K-12 Students,* (Alexandria), VA.

Tunmer, W. E., Pratt, C., Herriman, & M. H. (1984). *Metalinguistic Awareness in Children.* Berlin, Germany: Springer.

Williamson, J. & Hardman, F. (1995). Time for refilling the bath? A study of primary student-teachers' grammatical knowledge. *Language and Education, 9,* 117-134.

Winokur, S. (1976). *A primer of verbal behavior.* Englewood Cliffs, NJ: Prentice-Hall.

Wray, D. (1993). Student-teachers' knowledge and beliefs about language. In N. Bennett, & C. Carré (Eds.), Learning to Teach. London: Routledge.

Wright, T. (1991). Language awareness in teacher education programmes for non-native speakers. In C. James and P. Garrett. *Language Awareness in the Classroom.* Harlow: Longman.

Wright, T. (2002). Doing language awareness: Issues for language study in language teacher education. In H. Trappes-Lomax and G. Ferguson. *Language in Language Teacher Education.* Amsterdam: John Benjamins.

Zipke, M. (2008). Teaching metalinguistic awareness and reading comprehension with riddles. *The Reading Teacher, 62*(2), 128-137.

Zipke, M. (2011). First graders receiving instruction in homonym detection and meaning articulation: The effect of explicit metalinguistic awareness practice on beginning readers. *Reading Psychology*, *32*, 349-371.