

**DEVELOPMENT OF WRITTEN COMPLEXITY AND ACCURACY IN AN  
INTERMEDIATE TO ADVANCED GERMAN L2 SETTING USING  
WEIGHTED CLAUSE RATIO**

by  
**Gemini Fox**

**A Thesis**

*Submitted to the Faculty of Purdue University  
In Partial Fulfillment of the Requirements for the degree of*

**Master of Arts**



School Languages & Cultures

West Lafayette, Indiana

May 2019

**THE PURDUE UNIVERSITY GRADUATE SCHOOL  
STATEMENT OF COMMITTEE APPROVAL**

Dr. Colleen Neary-Sundquist, Chair  
School of Languages and Cultures

Dr. Alejandro Cuza Blanco  
School of Languages and Cultures

Dr. John Sundquist  
School of Languages and Cultures

**Approved by:**

Dr. Jennifer Williams  
Head of the Graduate Program

## TABLE OF CONTENTS

LIST OF TABLES .....	5
ABSTRACT.....	6
CHAPTER 1. INTRODUCTION .....	7
CHAPTER 2. LITERATURE REVIEW .....	11
2.1 Written Complexity .....	11
2.2 Writing Accuracy.....	14
2.3 Language Development & Study Abroad.....	16
2.4 Task Complexity, Resource & Tool Management .....	19
2.5 Weighted Clause Ratio .....	27
2.6 Chapter Conclusion.....	30
CHAPTER 3. METHODOLOGY.....	31
3.1 Participants.....	31
3.2 Samples .....	32
3.3 Data Analysis .....	34
3.3.1 Complexity Analysis .....	37
3.3.2 Accuracy Analysis.....	39
3.4 Chapter Conclusion.....	44
CHAPTER 4. RESULTS .....	45
4.1 Complexity Data .....	45
4.2 WCR Accuracy Data.....	47
4.3 Chapter Conclusion.....	50
CHAPTER 5. DISCUSSION.....	52
5.1 Review of Research Questions .....	52
5.2 Study Abroad and Language Development .....	55
5.3 Resource Management in Writing .....	58
CHAPTER 6. CONCLUSION.....	61
6.1 Limitations to the Current Study.....	61
6.2 Considerations for Future Study & Uses .....	63
REFERENCES .....	65

APPENDIX A: TASK TYPE EXAMPLES ..... 73  
APPENDIX B: DATA GRAPHS ..... 77

## LIST OF TABLES

Table 1: Table of Complexity Data .....	46
Table 2: Table of Accuracy Data .....	48
Table 3: Table of Error Counts .....	49

## ABSTRACT

Author: Fox, Gemini, T. MA

Institution: Purdue University

Degree Received: May 2019

Title: Development of Written Complexity and Accuracy in an Intermediate to Advanced L2 Setting Using Weighted Clause Ratio

Committee Chair: Colleen Neary-Sundquist

The primary focus of this study is to determine how clausal complexity and accuracy develop over the course of three academic years of intermediate to advanced-level German. This study aims to shed light on issues related to learner development of writing during advanced stages of language acquisition, particularly after conducting a study abroad. The main goal of this study will be to track the writing complexity and accuracy of multiple students longitudinally. This paper will identify Weighted Clause Ratio (Foster & Wigglesworth, 2016), as well as address Skills Acquisition Theory (DeKeyser, 2007), Interaction Hypothesis (Swain, 1985), and the Limited Attentional Capacity Theory (Skehan, 1998). In addition to this, the study will discuss the impact of a study abroad on the language-learning process, task complexity, and the language-learning plateau. Following a review of terminology, I will discuss how Weighted-Clause Ratio will be used to determine clausal accuracy and complexity. The data of this study will be analyzed with results shown in intervals throughout three academic years, comparing each of the three years with one another. Results indicate that accuracy increases drastically over the final two years when compared to the change in the first two years, confirming the effect that study abroad has on the written accuracy of learners, and the complexity showed improvements in some measures over the course of the study, but varied in other measures. I conclude the thesis by discussing by discussing the implications of these findings for our understanding of writing complexity and accuracy, and the long-term effects of study abroad.

## CHAPTER 1. INTRODUCTION

In this chapter, the usage of measures in SLA are discussed. The concepts discussed in the literature review are also briefly described in this chapter. The chapter concludes with the statement of research questions and hypotheses for the paper.

In SLA, measures are used to analyze the language development and performance of learners on various tasks. Examples of these tasks are an oral interview (Polat & Kim, 2013), or written instructions from a learner on how to handle a crisis (Gilabert, 2007). The mentioned measures are those of complexity, accuracy, and fluency (CAF measures). The three major components can be further broken down into individual measures as well. The choice of measure is important regarding the goals of a study, as well as proficiency level of the learner (Biber et. al., 2011). For example, researchers can choose to analyze the number of pauses in oral production as a measure of oral fluency (Chambers, 1997; Riggenbach, 1991), compare the mean number of T-Units to the mean number of error-free T-Units as a measure of written complexity (Iwashita, 2006), and count the number of error-free clauses in a text as a measure of written accuracy (Foster & Skehan, 1996).

The choice of measure can depend greatly not only on what is being studied, but also whom is being studied. Looking for advanced levels of subordination in beginner-level work will yield little-to-no data, whereas checking for the total numbers of errors in a text with advanced learners may yield results that are too general (Ortega, 2003). With the correct choice of measure and proper application, comparative data can lead to a richer understanding of learner development in SLA. One such population of interest in SLA, and for myself included, is the impact of language skills for students that have participated in a study abroad program for extended periods of time.

This group of learners are of interest, due to the immersion and exposure to their L2. This type of immersion is typically not found within normal learning contexts, aside from environments such as immersion schools. During a study abroad, learners are typically exposed to large amounts of the language and will likely have to interact in the language if they wish to succeed in the country. Due to the exposure and interaction with the language, many learners experience large amounts of language growth during these periods. This growth is typically in verbal communication, in which studies have found data supporting the positive impact of study abroad (Leonard & Shea, 2017; Wright, 2018). However, while many studies do have data on improvements of oral proficiency, there are very few studies conducted that have looked at improvements of written proficiency due to a study abroad (DeKeyser, 2007, Pérez-Vidal & Juan-Garau, 2009). Having additional studies regarding L2 writing proficiency after a study abroad will be beneficial to determining the overall impact of study abroad, as well as aid in understanding to L2 writing development.

The writing performance of intermediate to advanced-level German students over three years is investigated in this paper. In the first year, students have intermediate proficiency in German, and finish with advanced-level proficiency. The advanced proficiency is marked by students conducting a study abroad for six to twelve months. The measure used to analyze the data in this study is Weighted Clause Ratio (WCR), which compares errors made within a clause to the comprehensibility of the clause itself. I chose to use WCR because it analyzes accuracy with both a grammatical and communicative aspect (Foster & Wigglesworth, 2016; Evans et. al., 2014, Kuike & Vedder, 2007). This study is conducted in hopes to shed additional light on how a learner's writing complexity and accuracy may develop over the course of three academic years. In addition, the implications of study abroad on a learner's writing ability are also discussed in this

paper. I would like to discuss some main concepts that will be analyzed throughout this paper and influence the results of this study:

- Complexity and Accuracy as measures of L2 writing proficiency
- Development of learner skills and the “language-learning plateau”
- Impact of a study abroad
- Attributing factors of task complexity
- Weighted Clause Ratio as a measurement of written accuracy

The goal of this study is to determine how the written complexity and accuracy of learners develops longitudinally over three years. As the study is composed of two of the three major CAF measures, analyzing written complexity and accuracy can provide insight to the overall L2 writing proficiencies of the learners. These two measurements can be influenced by the study abroad experiences of the learners, as a study abroad can greatly impact language development (Pérez-Vidal & Juan-Garau, 2009; Carlson et al., 1991; Golonka, 2001; Yager, 1998). Determining if study abroad has any implications on writing performance will be beneficial for the field of SLA, as the impact of study abroad is already a largely studied field. Lastly, determining the impact of task complexity on written performance is also an area of interest, as studies have found varying results on the topic (Ishikawa, 2006; Gilabert, 2007; Nuevo, 2006).

The research questions examined in this study are identified below:

1. Based on the complexity measurements in this study, will the writing complexity of the students increase over the course of three years?
2. Based on WCR, will the writing accuracy of the students increase over the course of three years?

3. How will the average score ranges of writing complexity between the first two years (Year A & Year B) differ from the final two years (Year B & Year C)?
4. How will the average score ranges of writing accuracy between the first two years (Year A & Year B) differ from the final two years (Year B & Year C)?
5. Based on the measurements in this study, how do any changes in complexity and accuracy over the three-year timespan compare to one another?

These research questions encompass the learner development in both written accuracy and complexity over the three years. They serve to analyze the language progression for the years individually, as well as comparatively. The final question will compare how complexity and accuracy develop, to see whether there are any apparent increases or decreases for either measurements in relation to one another.

## CHAPTER 2. LITERATURE REVIEW

This chapter reviews the literature of the concepts present in Chapter 1. In sections 2.1 and 2.2, I will address issues in the literature regarding written complexity and accuracy. Section 2.3 and 2.4 focuses on previous studies of study abroad and task complexity. Lastly, Weighted Clause Ratio (WCR) is discussed, including its uses, benefits, and drawbacks.

### 2.1 Written Complexity

When making assessments on a learner's ability to produce their L2 in written contexts, it can be analyzed through fluency, syntactic and lexical complexity, and accuracy (Wolfe-Quintero, Inagaki, & Kim, 1998). The first of these two measures, complexity, can be analyzed in many ways. Complexity can be looked at lexically, regarding how complex, advanced, or unique the words in the produced text may be. Complexity can also be measured quantitatively, based on the number of words in every clause, sentence, and text, as well as the number of T-Units, clauses, and sentences in a text as well. Typically, in a classroom setting, rubrics that assess writing complexity will look for the types of words used, whether they are at the level of which the learners are, or the types of structures used and whether they or repetitive uses or if thoroughly expressed ideas.

Written complexity will first be subdivided into two categories: grammatical and lexical complexity. Grammatical complexity is composed of two primary parts, morphological complexity (Bulté & Housen, 2012; Brezina & Pallotti; 2016), and syntactic complexity (Hawkins, 2001; Lu, 2010; Crossley & McNamara, 2014), with morphological complexity concerning itself with the many inflections present in different languages, whereas syntactic complexity concerns

itself with various grammatical structures. Lexical complexity, however, pertains to how diverse the word usage is within a text, which should become more diverse as learners advance in their language skills (Crossley, Salsbury, & McNamara, 2009). Because these parts are constantly developing with the learner's language skills, the following definition covers written complexity and its constituents: "Grammatical... complexity mean that a wide variety of both basic and sophisticated structures and words are available to the learner" (Wolfe-Quintero, Inagaki, & Kim, 1998; 69, 101).

This definition is important, as it will help identify the complexity of structures used by learners as they progress in language development. Even at advanced levels, learners have basic structures to choose from, therefore syntactic complexity can vary. However, it is expected that learners will produce more complex structures and more sophisticated vocabulary as they improve their language skills. The effect of study abroad on syntactic complexity will be closely monitored in this study, to see how learners develop in their written complexity before and after the study abroad experience.

The ways in which complexity is measured must also be taken into consideration. In an article written as a guide to CAF measures, many different complexity measures are discussed, including their benefits and drawbacks (Norris & Ortega, 2009). The measures presented, although being used to measure syntactic complexity, measured syntactic complexity in various ways. For example, the mean length of a T-Unit or clause provide syntactic complexity that is "easy to calculate" (p. 559). Other methods, such as comparing subordination to T-Units in writing or the mean number of clauses per T-Unit, serve as additional syntactic complexity measures that fulfill their own specific purposes (Storch & Wigglesworth, 2007; Ellis & Barkhuizen, 2005).

Not all measures are intended to be used for all levels of writing, as mentioned in Ortega (2003). In Bergman and Abrahamsson (2004), a scale of syntactic features was created regarding learner proficiency levels. The scale began with beginner L2 learners and simple sentence structures, and advanced L2 learners were characterized as having sentences containing greater clausal diversity. Considering which measure to use for the proficiency of the writer is important to yield accurate results. A measure used to analyze a complex language feature may not be effective in beginner writing data, just as simple measures may yield results that are too general for a study.

L2 writing development is typically measured in four areas: fluency, accuracy, syntactic and lexical complexity. Nitta and Baba (2014) studied how task complexity affected syntactic complexity and fluency. Syntactic complexity was measured by finding the average sentence length and STRUT (sentence syntax similarity, all sentences across paragraphs), and lexical complexity was based on the lexical sophistication and diversity of the word choices within the text. This study addressed both aspects of the complexity definition used for the current paper, and the use of average sentence length to determine complexity will be helpful for this paper as well.

Biber et. al. (2011) took a different approach to complexity measures and determined how measures associated with spoken complexity would function in L1 academic writing contexts. Many measures were analyzed, and it was found that some spoken measures performed strongly in academic writing, and others did not. This was due to easily acquired grammatical features found in conversations were acquired in early stages, but complex structures occurred more frequently in adulthood. It is important to consider the performing environment when assessing complexity, as the choice of measure could greatly impact how complexity is assessed.

Using spoken measures in academic L2 writing, as hypothesized by Biber et. al., should return similar results. However, the writing proficiency levels and learning environments of the learners need to be taken into consideration. If a learner is an adult exposed to academic writing in the L1, but their L2 language skills are weak, the measures more suited for academic writing may be less effective measuring their writing complexity. These factors regarding the choice of complexity measures chosen for this study, as inappropriate measures would not yield substantial data.

## 2.2 Writing Accuracy

Writing accuracy, in a similar manner to that of writing complexity, has many ways in which it can be measured. An example of a written accuracy measure is the identification of every error in a text. Accuracy can be determined by other measures, such as the accuracy of a production unit (Crookes, 1990), morphology, or in the case of this paper, a comprehensibility-to-error clausal ratio (Foster & Wigglesworth, 2016).

However, accuracy has a definition that is more widely agreed upon than complexity. A common definition regarding accuracy is measuring how much a learner's performance differs from that of a native speaker (Hammerly, 1990; Pallotti, 2009; Wolfe-Quintero, Inagaki, & Kim, 1998). The closer a learner is to producing the language as a native speaker would, the more accurate they are. This applies in both a grammatical and communicative context, which will be useful in this study, because the learners will be assessed in how accurately they can convey their ideas in written production. The choice of accuracy measures, though, will be just as important as the choice of complexity measures.

In a study analyzing the validity and reliability of various accuracy studies, measures taken ranged widely in scope of analysis, from holistic scales and number of errors, to errors per a specific unit and error severity (Polio & Shea, 2014). Regarding reliability of measures, holistic measures tended to underperform measures of clausal representation. Holistic measures referred to measures of general accuracy, such as the number of errors present within a sample. Inversely, holistic measures held up more in respect to validity. Considering the findings from the study, it was determined that many accuracy measures were validated and cemented in reliability to assess written accuracy in the L2. However, it was also worth noting that a choice of accuracy measure is reliant based on the needs of the study. The current study requires a measure that can assess how accurate learners can be in written production. Therefore, WCR was chosen for this study.

Another area of interest for written accuracy is how the accuracy and errors progress within learner proficiency levels. In a study tracking specific error locations, learner data was taken from a corpus of English as a Foreign Language (EFL) learners ranging from B1-C2 on the Common European Framework of Reference (CEFR) (Thewissen, 2013). Errors chosen consisted of grammatical, lexical, punctuational, and even style errors, all of which were marked throughout the levels of learners. Interestingly, the results pointed to widely different development patterns, with some errors decreasing in frequency over time, whereas others would initially increase and then slowly regress. Although general improvement of accuracy did occur linearly from B1-C2, error frequency also decreased more in the B1-B2 proficiency levels and decreased less in B2-C1 and C1-C2. Development trends will also be important to analyze in this study, as the accuracy of the learners selected may not improve linearly regarding proficiency level throughout the three-year timespan.

### 2.3 Language Development & Study Abroad

There are a plethora of empirical data reinforcing the idea that both input and output in language development are necessary for acquisition in an L2 (Swain, 1985, 2000; Izumi, 2002; VanPatten, 1990). This idea is based on the Interaction Hypothesis, that learners must receive comprehensible input, understand it, and produce output accordingly (Long, 1996). The output produced, in a similar manner, should be comprehensible to other interlocutors. This is true for both written and spoken language production, and the dominant types of interaction can vary, based on the means which learners most frequently communicate.

This situation is particularly true for a study abroad, as learners are presented with vast opportunities to improve upon their oral production when studying in a foreign country (DeKeyser, 2007). There are many studies which have shown a strong effect regarding the impact that study abroad has on oral performance, especially in the realm of CAF measures (Carlson et al., 1991; Golonka, 2001; Yager, 1998; Allen, 2002; Davie, 1996; Freed 1995; 2004). However, there are very few studies having been conducted that look at written improvement over the course of a study abroad period. With what few there are, some show a positive effect of study abroad experiences, whereas others show zero to even negative effects that a study abroad experience could have on an L2 learner's writing proficiency (Freed, So, Lazar, 2003). This paper will prove that a study abroad is beneficial for a learner's writing proficiency.

Pérez-Vidal & Juan-Garau (2009) conducted a study which analyzed learner development in the L2 after a period of study abroad. The study spanned four periods, and the samples were analyzed for fluency, accuracy, and syntactic and lexical complexity. Their measure of fluency was used as words per clause and words per text, whereas accuracy was measured using errors per word, and grammatical and lexical complexity were used to measure written complexity. It was

found that learners improved on the written fluency and lexical complexity after the study abroad period, which can be attributed to the transfer of experience from oral interactions to learner's written work. However, in the other domains of measure, learners hardly showed signs of improvement, and even fell in performance, showcasing that learning was not linear throughout the study period, with a lag period in the middle of the study abroad (Thewissen, 2013). This also shows that certain aspects of performing may take precedence over another (Ortega, 2003). For the means of this study, it will be important to not only check learner developmental patterns throughout the three-year period, but also to check for any significant lag in results after the study abroad.

A similar study compared how students conducting a study abroad in L2 language use would differ than students who did not participate, both in oral and writing exercises (Sasaki, 2007). In this study, it was found that the students that participated in the study abroad improved in their L2 writing ability, especially regarding text length and increase in lexical complexity. Students who also participated in the study abroad were found to more likely to write better L2 compositions than those who did not participate. However, both groups improved on translating ideas from their L1 into their L2 in written production. These results indicate that both the study abroad group and the at-home group improved similarly, although not linearly, throughout the one-year study. Regarding the current study, learners should show similar growth in the realm of increased text length and increased complexity, if not lexically, then at least syntactically, due to the additional language exposure from the study abroad.

Even though a study abroad provides learners with the opportunity to immerse themselves in the target language, there are instances in which the language development may slow down or even plateau. This is known as the "language-learning plateau", which has typically been found

to occur between the intermediate and advanced levels of language learning (Richards, 2008; Shormani, 2013). Due to language learning happening at a quick pace, the learner may notice a period in which their skills do not improve as quickly as they once had at lower levels. This can be attributed to the number of rules left to learn in the target language at that stage, as well as advanced stages being periods in which learners refine their skills, not learn new ones. Although this plateau may be present in latter stages of learning, external exposure from the study abroad should push learners beyond the plateau, due to the refinement of language skills while abroad.

The language-learning plateau and studies conducted about the topic primarily concern themselves with EFL learners. One study analyzed this plateau, not only in respect to learners plateauing, but also how instructor influence may play a role in the language development process at intermediate-advanced levels (Mirzaei et. al., 2017). To accurately assess the impressions of the plateau, both language instructors and learners were selected as participants and provided with qualitative interviews, gauging their understanding of the subject. On the side of teacher-related variables, participants from both groups mentioned that lessons need to be diverse and engaging. However, language instructors made note that at higher levels, texts should be complex and demanding, which in turn caused some students to feel as though their skills were no longer improving, and they would understand fewer texts as they progressed with their language skills. From a learner perspective, some learners no longer had the same enthusiasm as they did when initially learning the language. This decline in motivation led to a declining drive to learn and master the language. Lastly, learners believed they were no longer progressing, noting, for example, “after three years, I still speak slowly and have trouble understanding at time”.

The factors mentioned above regarding instructors, learners, and problems that learners face are those that lead to the formation of the language-learning plateau. Yet, these factors do not

permanently restrain learners from advancing further in language development. This is stressed in Richards (2008), in which both issues and solutions to the plateau are addressed. The causes fall in line with those mentioned in Mirzaei et. al. (2017), who provided solutions that extend into both the realms of instructor and learner. For instructors, it is important that they do not overload learners with complex tasks that would cause learners to focus more on structure than meaning. However, if a task is to be issued that may prove difficult, learners should be provided with key vocabulary and explicit knowledge of structures to ensure that they stay motivated and interested in the task. Instructors should aid in pushing learners towards independent thinking and learning by providing them meaningful and current activities to give learners the opportunity to develop on their own.

Learners must diligently improve their language skills through communication with other learners and work with complex structures, despite being uncertain about those structures. Even regarding lexical items, learners will benefit more if they not only determine what the meaning of a word is, but also what they can group a word with (Nagy, 2005; Stanovich, 1986). The more learners can act independently with their language skills, the further they will develop past the language-learning plateau. This study will check for the present of a plateau during the three-year period. Based on the results, it will also assess if a plateau had been encountered and advanced from.

#### 2.4 Task Complexity, Resource & Tool Management

An important aspect of language production which must be taken into consideration is the task complexity that learners encounter when producing the language. Due to the variable nature that tasks consist of, knowing the demands of a task on learners can help better determine which

tasks will be more effective for a desired output, whether that be a communicative goal or analytic use of a target form. The findings on task complexity have been varied within SLA. Some have found that increases in task complexity often lead to increases in complexity and accuracy produced by the learners (Ishikawa, 2006; Kuiken & Vedder, 2007; Kuiken, Mos & Vedder, 2005). Others, however, have found mixed results due to increased task complexity (Gilabert, 2007; Revesz, 2011; Gilabert, Baron, & Llanes, 2009). Other studies indicated no significance at all (Nuevo, 2006; Lee, 2002). The mixed findings make it more pertinent to find concrete results within the current study.

In a study analyzing the variable effects of task complexity regarding oral and written production, learners were required to complete tasks of varying difficulty in their respective mode (Vasylets et. al., 2017). Learners were tasked with simple and complex forms of the Fire Chief Task (Gilabert, 2007), which tasks learners with handling a crisis through scenario planning. The simple version of this task involves a crisis that is much less serious in nature, requiring lower involvement with reasoning abilities (Révész et al., 2016). Regarding the assessment, learners were asked to produce orally and written for both the simple and complex tasks. It was found that, although oral production contained more ideas produced for both tasks, written production was syntactically more complex and accurate as well, with significant results found regarding time on task. Both speakers and writers were provided equal planning time. Due to the immediate output nature of oral production, speakers had no time to consult additional resources, either physical or mental. Writers, on the other hand, were not limited to the amount of time they spent writing. Speakers were confronted with the entirety of the task complexity immediately after planning, whereas writers could write, reflect, revise and continue writing their compositions. Due to this task on time, the task complexity was effectively reduced with respect to spoken output. The

written data here are beneficial for the current study, as learners produced written samples that were unrestricted regarding time on task. Also, they were provided texts for some samples, which should reflect a reduced task complexity.

The effect of task complexity is one that is commonly researched throughout the field of SLA, and for the purposes of this study, the impact of task complexity on written work in the L2 will be analyzed. In a study conducted measuring the effects of task complexity in L2 writing complexity, learners completed three writing tasks, each with increases in levels of complexity (Frear & Bitchener, 2015). Each of these tasks was designed to slowly increase the demand on the learner's cognitive attention; by the time of the third task, the impact of task type was apparent. The increases in task complexity were not solely limited to the level of questions posed to the learners, but also the amount of reasoning required on part of the learners to produce language for the task. When analyzing the results regarding task complexity, it was found that lexical complexity improved with task complexity, whereas syntactic complexity did not increase as task complexity rose. However, the tasks were carried out in a short time period. The current study, on the other hand, contains data from students over a period of three years. I believe that, despite Frear and Bitchener's results, the findings of the current study will show an increase to syntactic complexity, due to the period abroad.

This can be attributed to mental resource allocation when producing the language. However, results showing the inverse have also been found. Higher working memory can have a strong effect on learner's performance (Yi & Luo, 2013). The increase in task complexity could also be the result of the increased lexical complexity demonstrated by the learners. This is due to complex tasks often requiring the use of complex language. This sort of trade-off is reflected in

Peter Skehan's and Peter Robinson's assessment of working memory, called the Limited Attentional Capacity (Skehan, 1998) and the Cognition Hypothesis (Robinson, 1995) respectively.

In the Limited Attentional Capacity model, learners have one pool of finite attentional resources, from which they can draw from. Learners will focus primarily on one aspect of production, leaving less resources to devote to other processes. As task complexity increases the demand on attentional resources of learners, the learners will prioritize specific aspects of production less, resulting in a trade-off. Robinson's Cognition Hypothesis states that there are multiple pools from which learners can draw from, if the tasks do not encompass the same domain. For example, if a learner were writing, they would have to choose to prioritize the three CAF measures in any order, because they all are all aspects of production. An example of this trade-off is seen in Bergsleithner (2010), where learners showed greater improvement in complexity than accuracy. For the purposes of the current study, the varied task complexity will be taken into consideration when determining written complexity and accuracy as trade-offs for the learner's mental resource allocation.

As mentioned above, the allocation of mental resources, or "working memory capacity" (WMC) (Harrington & Sawyer, 1992), plays a large role in learners completing complex tasks: as the more complex the task is, the larger amount of attention is required for learners to perform. WMC equates to the learners short-term memory, as well as their ability to hold and process information. WMC can be split into two categories: storage and processing. Storage is used for chunking and encoding of lexical and semantic information, whereas processing occurs in the production of the language. It is important to consider WMC not only when learners process new information and perform, but also when learners interpret feedback, which will be seen in the study below.

In line with mental resource allocation and task complexity, stress from task conditions can greatly impact how well a learner performs on a certain task. Stress, defined by MacIntyre and Gardner (1994) as “the feeling of unease, worry, nervousness, and apprehension experienced in... a second language”, has many origins in the L2, as it can emerge from task conditions, as well as the learners’ needs and abilities to communicate effectively on a given subject. It is difficult for researchers to measure the immediate impact of stress on learners, as it can be closely attributed to task complexity. Due to the emotional impact brought about with stress, it is more effective for learners to be issued surveys, interviews, and hold discussions to gather information on how learners *felt* during a task. The impacts of stress in learner performance can be seen firsthand in the Fire Chief task used by Gilbert (2007) and Vasylets et. al. (2017), as the task complexity was “increased” by creating a larger crisis for the learners to work out. The increase in task complexity here was artificial; learners were required to work out the same task, within the same time, except the situation was changed to be direr. As learners were to treat this as a real scenario, they would need to choose the best course of action to handle the situation. The situational demands, which create stress for the learners, are what contributed to the difference in performance on that task in Vasylets et. al. (2017).

The increasing levels of complexity presented with tasks, which consequently increases the demand of learner resources, shows an increased stress induced by the task. In the same manner, learners with higher WMC were able to process the tasks more efficiently, which effectively results in lower amounts of stress. This supports the idea that stress demands additional resources from WMC: learners with higher WMC should ultimately perform better under higher amounts of stress than learners with lower WMC (Bergsleithner, 2010; Yi & Luo, 2013). This would support the findings above regarding lower WMC learners performing better than higher

WMC learners in tasks that had lower stress yields. Although unmeasurable due to lacking data in the corpus, these aspects will remain in consideration when analyzing the data in the current study.

To minimize stress, strategies will first be identified, then expanded upon. Learners can be provided a planning stage to prepare output, either written or spoken, which provides them time to solidify ideas in the language and improve performance (Foster & Skehan, 1996). As mentioned above, the environment in which learners conduct tasks may impact how stress affects the L2. For example, stress may increase due to learners presenting in front of their classmates or be minimized with informal interactions in small groups of other learners with their instructor. In conjunction with these methods, providing learners with resources while they are performing can also limit stress caused by a task. This should be both monitored and limited, however, since heavy reliance on said materials can lead to less authentic output and hinder language development (Kozlova & Presas, 2016). This is especially problematic for this study, as learners were often provided a text for later samples, and were either required to use it, or were permitted to use it as supplementary material.

A study analyzing how these strategies play a role in L2 writing stress was conducted with a group of EFL learners (Tsiriou et al., 2017). Writing anxiety in this article is defined as “writers who are competent enough to intellectually adhere to a task, but, nonetheless, face difficulty with the process of writing production” (Zheng, 2008). This is particularly true for foreign language writing, as learners must both be competent writers and skilled in the L2 to write with minimal stress. The article in mention addresses the idea of “cognitive apprenticeship”, which has learners follow an instructor or leader through the cognitive processes, instead of

conducting them themselves. This strategy greatly lowers the processing demand on behalf of the learner, in turn reducing the levels writing anxiety experienced during the task.

Grade-school learners were the basis of this study and were analyzed as to how well the mentioned facilitation efforts would reduce their stress. It was found that learners who were led through the cognitive processes before writing the tasks on their own performed with significantly less stress than the control group. Due to the lowered stress, learners were able to devote more attentional resources to the task, in turn allowing them to perform better as well. These results align with data from the previous studies mentioned above. In the same regard, the EFL data from this study can be translated over to learners of another foreign language, as the same mental facilities will be in use. Planning and load bearing of learners completing tasks will be an important factor for future studies analyzing L2 language production, especially in writing. Although learners in the current study were not provided with help from the instructor while writing, the texts provided to learners holds this functionality, to a lesser degree.

Another aspect of writing that needs to be taken into consideration is the learner's use of an internal monitor while writing. This "monitor", as mentioned in Stephen Krashen's Monitor Hypothesis, is one that learners may consult when they are provided the opportunity to reflect upon their output. Effectively, learners will consult prior knowledge before producing the output, which is not effective given the conditions of oral production. The usefulness of this monitor can be related to the learner's proficiency level.

This concept is reinforced in a study analyzing the development of the written languaging, or the reflection of one's own production language, of L2 English learners (Ishikawa, 2018). The study is metalinguistic in nature, and had students reflect on their language use by answering additional questions based on critical assessment of their own responses. Students of lower-

proficiency levels benefited more from the written languaging than higher-proficiency learners (Ishikawa, 2018). This study shows that, as competence in the language improves, less reliance on the internal monitor will occur. Similarly to Ishikawa's study, the learners in this study showed reduced reliance on monitor usage and written languaging due to an increased proficiency level, particularly in later samples.

A similar result is found in the study conducted by Sasaki, in which a group of EFL learners who studied abroad were compared to an additional group who stayed at home, all while learning the target language. Both groups, despite being low-intermediate in proficiency, had differing results regarding eagerness to write, as well as writing competence, despite both groups showing improvement in the language over the course of a year. Although both groups exhibited improvements in translating ideas from their L1 to the L2, only the study abroad group showed greater confidence and proficiency.

The use of a monitor in production of the L2 can also function as an inhibitor of acquisition, due to students storing explicitly learned information and referring to it during production (VanPatten & Oikkenon, 1996). The nature of a monitor requires that learners need to have larger periods of time to use it, making its use during speech impractical. However, explicit knowledge of the rule must be known to be recalled through the monitor. Once the learner has familiarized themselves well-enough with the rule, they can wane usage of the monitor. This strongly supports the results seen in Ishikawa, and somewhat supports the results of Sasaki, as lower proficiency learners will heavily rely on their monitor due to unfamiliarity with a language rule. Although learners in this study are writing and reading, the use of a monitor and provision of a text can limit the types of structures produced in the samples.

## 2.5 Weighted Clause Ratio

This paper concerns itself with the accuracy measure of Weighted Clause Ratio (WCR) (Foster & Wigglesworth, 2016). A component of this measure, ratio of clauses to T-Units, serves as a complexity measure. The ratio of clauses to T-Units serves as a complexity measure for the entire text: the higher this ratio is, the more complex the sample is. By taking each T-Unit and clause in the text into consideration, one can use clause ratios to provide a more accurate representation of the written complexity than with other measures, such as total word count (Norris & Ortega, 2009)

WCR measures written accuracy by considering clausal errors of varying degree. For every error in each clause within a text, WCR assigns a score. There are four weightings, or levels, of scores, and these are assigned based on the type of error present in the clause. The highest of the four levels is 1.00, with diminishing scores for the remaining levels based on the following criteria (Foster & Wigglesworth, 2016):

- *Entirely Accurate*: The clause is accurately constructed. A score weighting of 1.00 is applied for a clause at this level.
- *Level 1*: The clause only has minor errors (e.g. morphosyntax) that do not compromise the meaning. A score weighting of 0.80 is applied for a clause at this level.
- *Level 2*: The clause contains serious errors (e.g. verb tense, word choice, word order), but the meaning is recoverable, although not always obvious. A score weighting of 0.50 is applied for a clause at this level.
- *Level 3*: The clause has very serious errors that make the intended meaning far from obvious and only partly recoverable. A score weighting of 0.10 is applied for a clause at this level.

Each clause is analyzed for errors and assigned to a level, and the score weighting will be applied to that clause. Once all clauses have been scored, the scores of each clause will be added together and then divided by the total number of clauses, giving the WCR for the text. The final WCR will never exceed 1.00, as an entirely accurate text will have a maximum WCR of 1.00. The closer the WCR is to 1.00, the more accurately produced the text was. As an example, a WCR of 0.985 would be a highly accurate sample, whereas a 0.135 would be a highly inaccurate sample. It is important to note that, in a clause with multiple errors at varied levels, the more severe error will take priority. For example, if a clause contains both an error in morphosyntax and an error in word order, the clause will be attributed to Level 2, as the severe error takes precedence over the minor error.

WCR is the chosen measure for this study because, despite it being primarily an accuracy measurement, it is possible to measure complexity with it as well. It can determine the complexity for an entire text through the clause ratio. Moreover, it is a measurement that does not consider various types of grammatical features, unlike others that only focus on the amount of subordination in clauses, or verb positioning (Norris & Ortega, 2009; Ellis & Barkhuizen, 2005). Due to its general nature, this measure can be used both at earlier and later stages of proficiency to measure clausal complexity. However, it does not analyze lexical complexity, or the complexity of specific grammatical structures. Therefore, it will need to be supplemented to provide further insight for the complexity changes.

Unlike the clause ratios being a general measure, WCR as an accuracy measure is multidimensional. WCR not only focuses on the types of errors produced, but also how much those errors impede the comprehensibility of the clause. In this regard, a learner's accuracy and their ability to communicate effectively in written production are assessed at the same time. This

measure is also more comprehensive than other measures of accuracy, which can be as simple as checking the total number of errors made per sentence, or even overall in the text. While it does not focus on any specific type of grammatical error, which has been a drawback in some studies (Evans et. al., 2014), it does provide a more wide-ranging measurement of the development of written accuracy.

Some studies have pointed out some drawbacks to using WCR. Evans et. al. (2014), for example, pointed out that WCR as a measurement of accuracy was not very practical regarding time. Although requiring additional time to analyze a sample, the measure proved to hold up when compared to how well other measures perform, such as Error-Free Clause Ratio (EFCR) and Error-Free T-Unit Ratio (EFTR). Evans et. al. found that WCR takes at least twice as long to analyze a paragraph as other measures (p. 43). This is true when viewing whether WCR is a valid construct, as it determines how well a learner can accurately convey a meaning, and not how limited they are based on the errors they have produced. However, Foster (2016) used the measure to analyze L2 accuracy and found that Level 1 and 2 errors were the most disagreed upon, although these disagreements were resolved after fair amounts of discussion.

I have also worked with WCR firsthand when analyzing how written accuracy can develop over the course of an academic semester. I analyzed three samples of beginner-level L2 Italian, one of which was written, and two were spoken. The spoken samples were recorded and transcribed outside of class sessions, then analyzed for errors by a native speaker of Italian. The topics for the spoken samples pertained to the everyday life of the learner, and the learner spoke for 90-120 seconds for both samples. However, the one written sample was taken as part of an in-class writing assignment where the errors were marked by the instructor. The task pertained to material covered in-class, and the learner was restricted to twenty minutes of unprepared writing,

without the opportunity to consult additional resources. The purpose of the analysis of these samples was to determine how accurate the learner was in written and spoken contexts. Similar to Evans et. al. (2014), the samples took much longer to rate, yet they were able to be analyzed for communicative accuracy and sentence complexity. However, because I was the only rater and I laid out what types of errors belong to what level, I did not run into issues in distinguishing between a Level 1 or Level 2 error.

## 2.6 Chapter Conclusion

In this chapter, issues from previous literature were discussed. In sections 2.1 and 2.2, written complexity and accuracy were addressed, definitions were provided, and measures of the two constructs were explored. In section 2.3, the effects of study abroad and the language learning plateau were discussed, particularly regarding their impact over time, and to the current study. Section 2.4 discussed the issue of task complexity, and how attentional resources are influenced by task complexity to impact writing performance. Lastly, section 2.5 addressed WCR, the benefits and drawbacks of the measure, as well as previous experience using the measure.

## CHAPTER 3. METHODOLOGY

The methodology for this study is discussed in this chapter. The participants are first identified, followed by a short discussion on the samples gathered by CLEG13 (Corpus for Learners of German). After this discussion, the data analysis will follow, with examples from the corpus representing the error types presented by Foster & Wigglesworth (2016).

### 3.1 Participants

The data for this study were obtained through the CLEG13 corpus (Corpus for Learners of German), which exists under a subset of the larger corpus Falko, an open-source corpus found at <https://korpling.german.hu-berlin.de/falko-suche/>. Data were compiled in this corpus over three years. Data for this corpus were collected at Lancaster University from students in the German program at the university. The writing samples in the corpus are “expository-argumentative” texts in which learners were asked to think critically and discuss their own opinions regarding a specific topic. These types of texts were chosen from learners, as they are the type of texts that they would be expected to produce during their normal studies at Lancaster University.

The participants of the corpus followed the curriculum of Lancaster University, which had assessment processes in place for students entering Year A, as well as a study abroad between Years B and C. Each student was between 18 and 19 years of age at the start of their program. For students entering Year A, they were required to pass a “post-A level” examination at the university. The exam represents approximately five to seven years of German instruction at a secondary school, wherein they will begin in Year A at Lancaster University. CLEG13 stresses that, if looking at the three-year program through the lens of CEFR, Year A is the equivalent of a B1/B2 proficiency, Year B is a B2 proficiency, and Year C is a C1/C2 proficiency level. This

jump between Year B and Year C is present due to the students of the German program and Lancaster University attending a six to twelve-month study abroad before the third year of their degree program.

For the purposes of this study, I chose to follow fifteen students from CLEG13 longitudinally for three years throughout their program at Lancaster University. Students were selected if they had two samples in Years A and C, as well as one sample in Year B. It is possible that the duration of which the learners studied abroad in Germany between Year B and Year C could differ, from either six months to one year. This constraint is unfortunately unavoidable due to the limited data present in the corpus regarding duration of the study abroad period. The five samples gathered from each student were then analyzed with WCR. These samples from each student would provide a benchmark as they progress through the years, and prove sufficient when comparing between years as well, in particular Years B and C.

There are a few things to note about the participants in CLEG13. Although each participant conducted a study abroad, data on their language use abroad is not present in the corpus. This is problematic in determining how study abroad is affecting the learners' language development. Also, despite learners taking a proficiency test before starting their program, the specific score received by the learners is also not identified. Although the general proficiencies of the learners are known, exactly where they lie within a level is still ambiguous. These issues are further addressed in the limitations section of the conclusion.

### 3.2 Samples

The samples in the corpus were collected by researchers at Lancaster University. Samples in CLEG13 were collected uniformly from all participants, ranging from one sample every two

months over the course of ten months down to one sample every few weeks. As the participants moved onward throughout the years, data were still routinely collected and separated from previous years by a marker on the participant ID, as well as being tracked in three categories: Year A, Year B, and Year C. Data samples for the corpus participants vary in quantity, with some participants having two or three samples available for each academic year, and other participants having eight to ten samples submitted for an academic year. This variance can be seen in the later years of data collection, as more participants produced writing data for the corpus than in earlier years. For this study, 5 samples were gathered from 15 participants, resulting in 75 samples in total.

The samples for this study came from different tasks, but all involved a form of student opinion. This ensured genuine output on behalf of the learner and allowed for all samples to be similar. For example, a student could write an *Aufsatz* or a “general, argumentative essay”. Essays of this nature exhibit a range of topics, varying widely from Christmas experiences to violence in the media. However, for the data collected in this study, there are texts labeled as *Kritische Zusammenfassung*, a “critical summary”, or a *Kritische Kommentar*, a “critical commentary”. For the first of these two types of texts, students were asked to read a text and provide a short summary in German. They were also asked to take a critical stance on the texts and expand upon it, providing outside and personal information that might pertain to the text. An *Aufsatz* was typically written during the first two years of study.

The latter of the two critical-essay types is the critical commentary. These essays were conducted during the final year of study at the university, after the period abroad. In these essays, students are instead asked a question and required to write critically about the topic as much as possible, developing their own arguments and basing claims on their own knowledge. They were

provided a text that served as additional information to the topic yet were not required to read and relate to it while writing. Examples of each text type are present in the appendix. In order to maintain uniformity, 180 to 190 word segments were taken from each sample.

### 3.3 Data Analysis

Data analysis was conducted using WCR as a measure of development of written accuracy. Due to the multidimensional nature of WCR as an accuracy measure, it was not measured in tandem with other measures of accuracy. To measure complexity, I decided to supplement WCR with the number of simple sentences in a sample, and the average length of a T-Unit (Hunt, 1965; Ishikawa, 2007). The number of clauses at the respective accuracy levels and the weighted clause ratios themselves were kept individually, with that data serving as one-half of the overall data analysis process. The other half of the analysis involved averaging the data by year, to see how the accuracy and complexity measurements developed longitudinally over three years.

The samples collected from the corpus were marked for clause ratings by the author, a fluent speaker of German, in order to obtain an accurate measurement. Additionally, the other measures used in the study were marked by the author, ensuring reliability between the analysis of the two measures. The number of simple sentences was used as a contrastive tool to further flesh out the syntactic complexity of learners' compositions.

The expectation is that there is an increase in writing complexity and accuracy throughout the three-year timespan, as seen in Sasaki (2007). I expect the writing complexity to increase from the first year of data samples to the third year of samples. This is due to the anticipated language development of the learners. In a similar fashion to writing complexity, I also expect writing accuracy to similarly increase. The degree of these changes is unknown, particularly regarding

the impact of a study abroad. This is due to no control group being present in the study, therefore any conclusions drawn regarding the impact of study abroad are speculation. Language use while abroad was also not recorded. These issues are further addressed in the limitations of this study.

When comparing both written complexity and accuracy in groups consisting of Years A & B and Years B & C, I expect to see an increase in the later years' group. When analyzing the data for complexity, I expect the average score ranges for the first two years to be smaller in range and lower in value than the last two years. Inversely, I expect the average score ranges of the final two years to be larger, but higher in average overall score. I attribute the higher average scores to the generally beneficial nature that study abroad experiences have on the language learning process (Carlson et. al., 1991; Golonka, 2001). However, the impact that a study abroad can have on a learner can widely vary. Therefore, a wider range within the group is expected (Kitiabi, 2018).

I also expect the written accuracy of the learners to increase as they progress throughout the years of their studies, in a similar manner to how the writing complexity is expected to increase. Due to the progression from intermediate to advanced-learning stages, I expect the written accuracy to be lower in value for the first two years, with score ranges to be smaller in size as well. I expect the writing accuracy to show greater improvements during the last two years. This is once more in relation to the general benefits of study abroad observed by other studies. Again, due to the lack of a control group, the impact of study abroad in this study are speculation. Although the study abroad is taken into consideration, direct conclusions cannot be drawn from the experience alone.

Lastly, when looking at the complexity and accuracy data for any patterns, I anticipate that the writing complexity and accuracy will both increase with each another, but to varying degrees. I expect the increase to be apparent between the second and third years of student writing. I

attribute this to continued practice of complex language structures. As learners use increasingly complex structures, the opportunity for additional errors rises, whether they be syntactical, lexical, morphological. However, as the learners progress into advanced stages of learning the language, these complex structures should become more commonplace. It is here where written accuracy should see improvement as well.

Written accuracy and complexity are expected to improve over these three years, although the degree of the changes is difficult to predict. As mentioned before, there are few studies that have analyzed the impact of study abroad on written proficiency (DeKeyser, 2007). Basing judgement from these studies, it is expected that the written accuracy and complexity increase. However, the impact of the study abroad is not possible to assess without a control group. Therefore, the study abroad may play a factor in improvement and is considered alongside language development when analyzing data.

Due to WCR functioning as a measure of communicative accuracy, it can be used to identify grammatical errors and categorize them in severity. I chose this measure for this reason, as learners typically use a language to communicate with one another and attempt to make themselves understood as best as possible. The measure proves particularly beneficial in determining how well the learner's communicative accuracy improves over the course of three years. However, while communicative accuracy is assessed, the frequency of specific grammatical errors is not identified, which serves as a set-back for this measure. Although WCR is an accuracy measure, the clause ratio for each sample serves as a measure of complexity. To further supplement this complexity aspect, I also chose to record the average number of simple sentences present in the data. Samples indicate an increase in complexity when they contain fewer simple sentences.

### 3.3.1 Complexity Analysis

Each sample was first analyzed for the number of T-Units and clauses in the sample. T-Units are defined as a main clause and all possible dependent and independent clauses associated with the main clause. An example of a T-Unit is provided below (Learner ID: C1008\_04):

*Es gibt Menschen, die die gerne verbieten würden, obwohl sie sehr populär unter Kindern sind.*

There are people, who them gladly prohibit would, although they very popular under children are.

‘There are people who would gladly prohibit them, although they are very popular with children.’

The entire sentence above is considered a single T-Unit. It is composed of three clauses, because each clause contains its own subject noun and verb pairing. After T-Units were analyzed, clauses were counted. Clauses are defined as a sentence constituent containing a noun and verb pairing. Due to this, some of the samples within Year B had a high clause count, as sentences containing the “*zu- Infinitive*” were very common in these samples. Below is an example of the “*zu- Infinitive*” (Learner ID: B1001\_03):

*Mit mehr Geld haben wir auch bessere Gelegenheit, uns zu amüsieren.*

With more money have we also better chance ourselves to enjoy.

‘With more money, we have better opportunities to enjoy ourselves.’

In the example above, “*uns zu amüsieren*” is the “*zu- Infinitiv*”. Due to this clause lacking a grammatical subject, the use of the pronoun “*zu*” is required here. Thus, the sentence above contains two clauses: the main clause of “*Mit mehr Geld haben wir auch bessere Gelegenheit...*”, and the dependent clause of “*...uns zu amüsieren*”.

The average length of T-Units was determined by finding the total number of words in the sample and dividing that by the number of T-Units present. For example, a text containing 180 words and nine T-Units has an average T-Unit length of 20 words per T-Unit. Samples with higher average T-Unit lengths are more demanding on the learner's mental resources, thus showing signs of increased text complexity. This falls in line with Ortega (2003), in which average length of sentences was used as a measure of syntactic complexity.

The number of simple sentences present in the samples also represents text complexity. A simple sentence is defined as "A sentence with one independent clause and no dependent clauses" (Purdue OWL). An example of a simple sentence is provided below (Learner ID: C1005\_07):

*Einigen Akademikern stimmen die Ergebnisse nicht zu.*

Some academics agree with the results not.

'Some academics do not agree with the results.'

The above simple sentence matches the definition, in that it is a single independent clause, with no dependent clauses. For this study, the fewer simple sentences present in a sample, the more complex the sample is. This is due to the sentences experiences greater clausal diversity. This measure will change inversely to the clause ratio, which is discussed below.

The clause ratio for each sample was found. This was found by dividing the number of clauses over the number of T-Units present. This was insightful, regardless of the word length of the sample, as a sample containing 30 T-Units and 60 clauses will have the same clause ratio as a sample containing 15 T-Units and 30 clauses. The higher the clause ratio, the more complex each T-Unit was, due to the amount of subordination and coordination found within the T-Unit. Based on changes in the clause ratio, developments in complexity over time can be determined.

### 3.3.2 Accuracy Analysis

Regarding the accuracy data, samples were first analyzed by finding the number of fully correct T-Units, and from there, the number of fully correct clauses. It was done as so, as any correct T-Units would contain completely correct clauses, and T-Units with an error may still have multiple clauses that are still correct. An example of a fully correct T-Unit is provided below (Learner ID: C1005\_07):

*Der Artikel handelt sich um das Problem von Plagiat bei den irischen Universitäten.*

The article handles itself about the problem of plagiarism at the Irish universities.

‘The article is about the problem of plagiarism at Irish universities’

The T-Unit above consists of one sentence and one clause. Because no error is present, the above example would count as one fully-correct T-Unit and one fully-correct clause. Although the number of fully correct T-Units was not used in the final calculation to determine the WCR of a sample, they still have some use regarding the longitudinal view of each student regarding how well they can communicate entire ideas without an error. The same applies to clauses, holding even more value than T-Units, as something as simple as a spelling error can mark an entire T-Unit as not being fully correct, whereas it would only remove one clause of potentially many others from the final fully-correct clause count.

Clauses that contained errors were identified and categorized, with each error type landing the clause within a distinctive level. I have provided examples from the data of each error type at the various levels below. Level 1 errors hardly impeded the communicative value of the clause and required little to no extra effort to interpret the meaning due to the error. Errors that pertained to this category included errors in number (singular, plural), case and gender in definite and indefinite articles, as well as misspellings of a word, including noun capitalizations. Examples of each can be found below:

1. *Number* (Learner ID: A1005\_01)

*Ich bin ganz glücklich, dass es gibt vier meines Freundes Lancaster.*

I am very lucky, that there are four of my friend Lancaster.

‘I am very lucky, that there are four of my friend from Lancaster.’

2. *Case and Gender* (Learner ID: C1008\_04)

*... wenn ihre Kinder von gewalttätige Computerspiele fasziniert sind*

... if their children from violent computer games fascinated are

‘... if their children are fascinated by violent video games’

3. *Misspellings* (Learner ID: B1001\_03)

*Mit mehr Geld haben wir auch bessere Gelegenheit, uns zu amusieren*

With more money have we also better chance ourselves to enjoy.

‘With more money, we have better opportunities to enjoy ourselves.’

Each error of this level was weighted at 0.80. These errors were more common in the earlier samples than in later ones, as by that time, learners typically have a concrete understanding of number, case, and gender. The last subset of Level 1, misspellings, falls in a gray area regarding errors: the learner would not be able to misspell a word if they are speaking instead of writing. Although Foster & Wigglesworth (2016) do not make mention of this within their discussion of WCR, Kuiken & Vedder (2007) classify spelling as a “first-degree error” (p. 117). This lines up with the error levels present in the WCR analysis. For future studies, this could be changed to have misspellings omitted from the data analysis, as it could be assumed that learners know the word. This holds true of studies regarding spoken production.

Level 2 clauses are assigned a weighting of 0.50. These types of errors moderately impacted the communicative value of the clause, but the meaning was still able to be recovered

after the clause was reread in context with the rest of the sentence. Errors that would place a clause within this level include incorrect verb placement, verb conjugation and tense, a conjunction error, as well as an error in preposition usage. Examples of these types of errors can be found below:

1. *Verb Placement* (Learner ID: A1005\_01)

*Ich bin ganz glücklich, dass es gibt vier meines Freundes Lancaster.*

I am very lucky, that there are four of my friend Lancaster.

‘I am very lucky, that there are four of my friend from Lancaster.’

2. *Verb Conjugation & Tense*

a. *Verb Conjugation* (Learner ID: A1006\_01)

*Das ist einer der Sache, der ich bei Uni die meisten vermissen.*

That is one the things, that I at Uni the most miss.

‘That is one of the things, that I miss the most at the university.’

b. *Verb Tense* (Learner ID: A1006\_01)

*Normalerweise, würde ich jeden Wochenende ins Stadion gegangen.*

Normally would I each weekend to the stadium went.

‘Normally I would went to the stadium each weekend.’

3. *Conjunction Error* (Learner ID: B1005\_05)

*Es gibt auch Hausarbeit, aber ich denke (dass) es gibt ein Teil des Lebens.*

There is also housework, but I think (that) there is a part of life.

‘There is also housework, but I think (that) there is a part of life.’

4. *Preposition Error* (Learner ID: A1015\_01)

*In den Weihnachtsferien bin ich mit den Zug zu Hause in Hertfordshire gegangen.*

In the Christmas break had I with the train at home in Hertfordshire went.

‘During Christmas break, I went home by train to Hertfordshire.’

Errors of these types are errors that learners begin to identify in the earlier stages of language development, and by the advanced language levels, these types of errors are smaller in quantity. These errors were common throughout the earlier samples, but not as prevalent in later samples, in which Level 1 errors primarily dominated. However, due to the more complex occurrences of these errors, learners were expected to make these types of errors on occasion in the samples in Year C, especially errors regarding a missing preposition with the “*zu- Infinitive*”.

Lastly, Level 3 errors were errors that had the potential to completely hinder comprehensibility, as key components of clauses were missing. Only after much interpretation could some form of the original meaning be recovered. Therefore, clauses that landed within this level were attributed with the weighting of 0.10, having a major impact on the WCR of the learner. Errors that would bring a clause to this level include missing verbs and incorrect subject or pronoun usage. It is worth mentioning here that, in German, the subject and verb are always required in the sentence, despite the verb conjugation showing the subject of the sentence. Examples of the mentioned error types can be found below:

1. *Missing/Excess Verbs* (Learner ID: B1003\_05)

... *weil es mehr zum Leben als arbeiten (gibt) und Geld würde keine Rolle spielen*

... because there more to life than work (is) and money would not a role play

‘... because there (is) more to life than work and money would not play a role’

2. *Incorrect Subject/Pronoun Usage* (Learner ID: A1004\_01)

... *aber ich habe mir trotzdem Spaß gemacht*

... but I had me anyways fun did

‘... but I had fun anyways’

Errors of this level were rare, but one occurrence of these types of errors had a major impact to the WCR of the entire sample, as one Level 3 error equates to eight Level 1 errors. These errors also occurred more frequently in earlier samples than in later ones.

After the level analyses were completed, the scores were compiled and then divided over the total number of clauses, to determine the WCR. As a measure of accuracy, WCR offers insight into how the written accuracy of learners improved over the course of three years.

Learners improved greatly over those three years, especially after the learners went on a study abroad. To show the coding process for WCR, I have included a paragraph below following the guidelines listed above. Errors are underlined and marked with the respective error level. Each clause is separated by brackets, with a number following the brackets to indicate the error level of the clause. Below is the coded sample (Learner ID: A1009\_01):

[Die zweite Woche habe ich ein bisschen hausaufgaben<sup>1</sup> gemacht<sup>2</sup> bevor<sup>2</sup> das erster Weihnachtsfeiertag<sup>1</sup>.]<sup>2</sup> [Am erster Weihnachtsfeiertag<sup>1</sup> bin ich mit meinem Vater zu Hause geblieben,]<sup>1</sup> [weil meiner Mutter im Krankenhaus arbeiten musste.]<sup>0</sup> [Es war sehr ruhig aber entspannend.]<sup>0</sup> [Die dritte Woche habe ich mit meinem Freund verbrachte<sup>2</sup>,]<sup>2</sup> [weil er die vorherige Woche mit seiner Familie verbracht hat.]<sup>0</sup> [Also, habe ich das Neujahr mit ihm, meiner Eltern und meine Freunden verbracht.]<sup>0</sup> [Das Neujahr sind wir zum Kneiper<sup>1</sup> gegangen,]<sup>1</sup> [und spater<sup>1</sup>, (sind wir)<sup>3</sup> nach Hause zurück(gekommen/gegangen)<sup>3</sup>]<sup>3</sup> [um eine Party zu halten.]<sup>0</sup>

In the example above, there are a total of ten clauses. Of the ten clauses, five were fully correct. Two clauses were at Level 1, two clauses were at Level 2, and one clause was at Level 3. Scores are now assigned to each clause in this order: Fully-correct clauses are 1.00, Level 1 are at 0.80, Level 2 are at 0.50, and Level 3 are at 0.10. The raw scores would add up as so: 5.00 + 1.6

+ 1.0 + 0.1 = 7.7. Lastly, this is divided by the total number clauses to give us the WCR, which for the example above, is 77.00.

### 3.4 Chapter Conclusion

In this chapter, the methodology was first discussed. Participants from the CLEG13 corpus were used for this study, and 75 samples in total were collected for analysis. The analysis process for both complexity and accuracy was discussed, with examples of measures provided from CLEG13 student samples. Lastly, an example of the coding process was given, along with how the WCR is calculated from the coded data. The process above reflects the process used to calculate the WCR from the samples in the data chapter.

## CHAPTER 4. RESULTS

The results of this study are covered in this chapter. The complexity data is analyzed first, and the accuracy results follow. The results indicate that complexity fell for some measures, whereas it increased for others. The accuracy results show improvement over the three-year period. The means, standard deviations, and effect sizes are all compiled in tables within their respective sections.

### 4.1 Complexity Data

Each of the complexity measures have a varying impact on the written complexity of the learners throughout the study. The complexity data for this study will be compiled into Table 1, found below. Starting with the average number of T-Units found within a sample, the average number steadily decreased over the three-year period (*Year A:  $m=14.03$ ,  $sdv=3.05$ ; Year B:  $m=12.27$ ,  $sdv=3.01$ ; Year C:  $m=11.1$ ,  $sdv=5.00$ ). Although the average number of T-Units typically indicates changes of fluency, when compared to changes in clause count, can also be used to indicate complexity. The decrease in the average number of T-Units is also reflected in the average number of clauses present.*

Unlike the steady decrease present in the average number of T-Units produced each year, the average number of clauses produced indicated a much stronger change between the second and third years of the study (*Year A:  $m=27.43$ ,  $sdv=6.17$ ; Year B:  $m=27.27$ ,  $sdv=4.13$ ; Year C:  $m=24.5$ ,  $sdv=12.49$ ). This decrease in clause count suggests a decrease in clausal complexity within the sentences. When comparing both average number of T-Units and average number of*

clauses, however, average clauses decrease less than T-Units. This data represents the findings of the Clause Ratio, which will be discussed further below.

Simple sentences were also analyzed. As the number of simple sentences within the samples decreased, each sentence consisted of more clauses. The average number of simple

Table 1: Table of Complexity Data

	Year A <i>mean &amp; sdv</i>	Year B <i>mean &amp; sdv</i>	Year C <i>mean &amp; sdv</i>	Year A-B <i>Effect Size</i>	Year B-C <i>Effect Size</i>	Year A-C <i>Effect Size</i>
Average # T-Units	<i>M=14.03</i> <i>Sdv=3.05</i>	<i>M=12.27</i> <i>Sdv=3.01</i>	<i>M=11.1</i> <i>Sdv=5.00</i>	<i>d=0.58</i>	<i>d=0.28</i>	<i>d=0.71</i>
Average # Clauses	<i>M=27.43</i> <i>Sdv=6.17</i>	<i>M=27.27</i> <i>Sdv=4.13</i>	<i>M=24.5</i> <i>Sdv=12.49</i>	<i>d=0.03</i>	<i>d=0.30</i>	<i>d=0.30</i>
Simple Sentences	<i>M=3.97</i> <i>Sdv=2.31</i>	<i>M=2.47</i> <i>Sdv=1.92</i>	<i>M=3.00</i> <i>Sdv=1.49</i>	<i>d=0.71</i>	<i>d=0.31</i>	<i>d=0.50</i>
Average Word Count	<i>M=15.14</i> <i>Sdv=2.31</i>	<i>M=15.89</i> <i>Sdv=3.63</i>	<i>M=16.61</i> <i>Sdv=3.19</i>	<i>d=0.25</i>	<i>d=0.21</i>	<i>d=0.53</i>
Average Clause Ratio	<i>M=2.10</i> <i>Sdv=0.24</i>	<i>M=2.33</i> <i>Sdv=0.47</i>	<i>M=2.21</i> <i>Sdv=0.41</i>	<i>d=0.62</i>	<i>d=0.27</i>	<i>d=0.33</i>

sentences and the average clause ratio, which will be addressed later, show opposite patterns with one another. As the clause ratio increases with the presence of clauses, the number of simple sentences decrease (*Year A: m=3.97, sdv=2.28; Year B: m=2.47, sdv=1.92; Year C: m=3, sdv=1.49*). This indicates an increase in written complexity, as the sentences produced contain more clausal complexity. However, the samples produced during Year C were less complex with this measure than Year B. Analysis of task types and grammar in these samples will be addressed in the discussion section in order to explain why these results are present in the data.

The next measure of complexity, average words per T-Unit, also shows a consistent, although small, progression of written complexity by students. The complexity of the samples increased as the average length of the T-Unit increased. This is due to longer sentences requiring more attentional resources from the learner (*Year A: m=15.14, sdv=2.31; Year B: m=15.89, sdv=3.63; Year C: m=16.61, sdv=3.19*). These data are not only important on their own, but also

when reanalyzing the number of T-Units and clauses in the texts. When looking at that data, we can see that despite the average number of T-Units and clauses decreasing, the average length of a T-Unit is increasing. Although fewer T-Units and clauses are being produced, they are more complex due to their increased length.

Lastly, the clause ratio indicates the number of clauses for every T-Unit in the sample. This was calculated by dividing the total number of clauses over the total number of T-Units. For example, a sample with 10 clauses and five T-Units has a clause ratio of 2.00. The sample complexity increases as the clause ratio increases. An increasing clause ratio indicates an increase in complexity, because more clauses are present per T-Unit than in previous samples (*Year A:  $m=2.10$ ,  $sdv=0.24$ ; Year B:  $m=2.33$ ,  $sdv=0.47$ ; Year C:  $m=2.21$ ,  $sdv=0.41$* ). As previously mentioned, the data on the clause ratios and simple sentences have opposite patterns with one another. Just as with simple sentences, task types and grammar will be analyzed to explain the changes present with the data in Year B.

#### 4.2 WCR Accuracy Data

The accuracy data from the samples will be analyzed in this section, and the data were compiled into Table 2. Similar to the complexity data, it can be helpful to gauge how accurate a work will be from the number of fully correct T-Units, as well as the number of fully correct clauses. It is worth mentioning that the results of these components will only serve as supplemental information, and not be used for the final analysis for written accuracy. This section of WCR provides a general overview to the development of accuracy over the three years.

Table 2: Table of Accuracy Data

	Year A <i>mean &amp; sdv</i>	Year B <i>mean &amp; sdv</i>	Year C <i>mean &amp; sdv</i>	Year A-B Effect Size	Year B-C Effect Size	Year A-C Effect Size
Average # Fully-Correct T- Units	<i>M=3.93</i> <i>Sdv=2.94</i>	<i>M=3.33</i> <i>Sdv=2.35</i>	<i>M=5.03</i> <i>Sdv=2.08</i>	<i>d=0.23</i>	<i>d=0.77</i>	<i>d=0.43</i>
Average # Fully-Correct Clauses	<i>M=14.13</i> <i>Sdv=6.87</i>	<i>M=15.47</i> <i>Sdv=4.24</i>	<i>M=16.67</i> <i>Sdv=8.83</i>	<i>d=0.23</i>	<i>d=0.17</i>	<i>d=0.32</i>
Average WCR Scores	<i>M=79.95</i> <i>Sdv=7.59</i>	<i>M=80.35</i> <i>Sdv=3.91</i>	<i>M=90.07</i> <i>S=2.64</i>	<i>d=0.07</i>	<i>d=2.91</i>	<i>d=1.78</i>

(Year A:  $m=3.93$ ,  $sdv=2.93$ ; Year B:  $m=3.33$ ,  $sdv=2.35$ ; Year C:  $m=5.03$ ,  $sdv=2.08$ ). The increase in fully correct T-Units represents an increase in accuracy over the three years.

However, incorrect T-Units do not account for the fully correct clauses. The average number of fully correct clauses increased during the three years of the study (Year A:  $m=14.13$ ,  $sdv=6.87$ ; Year B:  $m=15.47$ ,  $sdv=4.24$ ; Year C:  $m=16.67$ ,  $sdv=8.84$ ). The results from the earlier years differ greatly from those found with fully correct T-Units. This is due to the strict nature of the previous measure: a T-Unit would be marked incorrect due to one incorrect clause, although other error-free clauses would be present. This measure allows us to get a better idea as to how learners use their language skills, yet it still only serves as supplemental information to the final WCR scores.

Continuing through the process of WCR, the number of errors at different levels also indicates the levels of accuracy which develop among each writer. Unlike the previous measurements, I found the total number of each error level for each individual year to determine what errors were more commonly produced. This corroborates with the WCR scores that will appear after this measure. Level 1 errors occurred the most in the first year, occurring more than

Level 2 errors by only 14 errors. Of all 399 errors in the Year A data samples, only 35 Level 3 errors occurred. In Year B, the most common error type was Level 2 errors, with 92 errors. Level 1 errors followed at 65 errors. The fewest number of errors was Level 3, with 24 errors. Lastly, Year C contained the fewest overall errors, yet still contained 154 Level 1 errors. This was the largest of any error type in this year, double than the number of Level 2 errors. Only seven Level 3 errors were in the Year C data samples.

Table 3: Table of Error Counts

	Year A	Year B	Year C	Year A-B <i>Effect Size</i>	Year B-C <i>Effect Size</i>	Year A-C <i>Effect Size</i>
Total Error Count	399	181	233	-	-	-
Level 1 ( <i>n, mean, sdv</i> )	<i>N=189</i> <i>M=12.60</i> <i>Sdv=4.84</i>	<i>N=65</i> <i>M=4.33</i> <i>Sdv=2.19</i>	<i>N=154</i> <i>M=10.27</i> <i>Sdv=8.76</i>	<i>g=1.91</i>	<i>g=0.80</i>	<i>g=0.34</i>
Level 2 ( <i>n, mean, sdv</i> )	<i>N=175</i> <i>M=11.67</i> <i>Sdv=4.15</i>	<i>N=92</i> <i>M=6.13</i> <i>Sdv=3.02</i>	<i>N=72</i> <i>M=4.8</i> <i>Sdv=3.97</i>	<i>g=1.46</i>	<i>g=0.38</i>	<i>g=1.67</i>
Level 3 ( <i>n, mean, sdv</i> )	<i>N=35</i> <i>M=2.33</i> <i>Sdv=1.88</i>	<i>N=24</i> <i>M=1.6</i> <i>Sdv=1.18</i>	<i>N=7</i> <i>M=0.47</i> <i>Sdv=1.06</i>	<i>g=0.45</i>	<i>g=0.98</i>	<i>g=1.04</i>

These results indicate that learners do improve in written accuracy over time, especially after returning from a study abroad. The number of higher-level errors in the Year C samples greatly decreased from the previous years. The number of Level 3 errors in Year C was particularly small. This shows improvement in the learner's ability accurately produce written work. The errors which greatly impeded communication in the earlier samples were not often present in the Year C samples. Some errors, for example spelling, would not occur in contexts outside of written work, which would make these error counts much lower if spoken data were to be analyzed. These error quantities corroborate with the WCR in the next section.

Lastly, the raw scores from the weighted levels and full clauses were added together and divided by the total clauses present to obtain the WCR for each sample. An increasing WCR represented a rise in increased written accuracy. WCR increased throughout the study, showing a much larger increase during the final two years (*Year A:  $m=79.95$ ,  $sdv=7.59$ ; Year B:  $m=80.35$ ,  $sdv=3.91$ ; Year C:  $m=90.07$ ,  $sdv=2.64$ ). The major increase in WCR between the last two years in comparison to the first two supports the idea that the study abroad played a major role in improving the learner's writing accuracy. When paired with the data on fully correct T-Units, clauses, and error counts, we can see this development further reflected in which in these fields continued to improve as learners progressed through the study.*

### 4.3 Chapter Conclusion

The results of this study were discussed in this chapter. First, written complexity was analyzed by the number of T-Units, clauses, simple sentences, the average words per T-Unit, and the clause ratios of the samples. This data indicated increases and decreases to complexity, depending on the measure in question. If complexity were based on the number of T-Units and clauses, the complexity over time would have decreased. However, it is necessary to consider the other complexity measures, as they showed improvement during the study.

Weighted Clause Ratio, a communicative accuracy measure proposed by Foster & Wigglesworth, was used to analyze clauses for their grammatical accuracy. The errors were categorized based on how much they impacted the communicative value of the clause. Minor errors were categorized as Level 1, moderate errors were categorized as Level 2, and major errors which halted the communicative value of the clause were Level 3. This is further explained in the

methodology. The accuracy data showed consistent improvement throughout the study, with major improvement occurring during the last two years.

## CHAPTER 5. DISCUSSION

In this chapter, the research questions are answered based on the results from the previous chapter. The results are then used to discuss the implications that study abroad has on written accuracy and complexity. Lastly, the effects of task complexity and resource management are addressed on how they impact the study.

### 5.1 Review of Research Questions

Here, I would like to address the research questions, starting with RQ 1, “Based on the complexity measurements in this study, will the writing complexity of the students increase over the course of three years?”. As the results indicate, the complexity measures vary. When looking at the average clause ratios, average length of T-Units and number of simple sentences, written complexity did indeed increase. Although the number of clauses is typically a measure of fluency, it can also be an indicator of complexity when analyzed with the number of T-Units. The number of clauses can allude to changes in coordination and subordination within a sample. Therefore, I would classify both number of clauses and number of T-Units as components to the complexity measure of clauses per T-Unit, or Clause Ratio.

The simple sentences, however, did indicate an improvement in syntactic complexity. This was particularly apparent when analyzed with the clause ratios. If a sample contains fewer simple sentences, then the sentences within contain greater amounts of coordination and subordination. Lastly, the average length of T-Units shows us that, despite the samples having fewer T-Units as the study progressed, they increased in length as time went on. This represents an increase in

complexity. Therefore, checking how the measures interact with one another serves a much greater purpose than if each measure were used on its own.

RQ 2, “Based on WCR, will the writing accuracy of the students increase over the course of three years?”. There is a clear increase over the period of the study in most all areas measured. Fully-correct T-Units and clauses showed an almost-consistent increase over the three years, and the WCR scores themselves also continually increased as well. The error counts, particularly Level 1 and 2 errors, fluctuated throughout the study. This change is observed in the samples between Years A and B, as Level 2 errors overtook Level 1 errors in Year B. As mentioned in the methodology, if a clause contained both a Level 1 and Level 2 error, the clause would be marked as Level 2. This was due to the more severe error taking priority. With errors in verb placement being the most common error in Year B samples, it is understandable that Level 2 errors were more prominent than Level 1 errors.

RQ 3, “How will the average score ranges of writing complexity between the first two years (Year A & Year B) differ from the final two years (Year B & Year C)?”. For the average number of T-Units and clauses, the complexity decreased in each of the three years. The average number of simple sentences and clause ratios decreased in complexity during the first two years, and then increased in complexity in the latter two years. Although clause ratio and simple sentences provided better representations of written complexity, the structures in the samples highly influence this, as the use of the *zu-Infinitive* greatly increased the clause ratios of Year B. This is because each *zu-Infinitive* indicated a dependent clause.

Only one complexity measure, average length of a T-Unit, had a positive trend throughout the entire three-year period, albeit a smaller increase during the last two years than the first two years. Although the T-Units may have been longer in the final year of the study, the samples had

smaller clause ratios and contained more simple sentences. If the learners had one more year of data, then the long-term effects on complexity could be seen. This would be useful in determining the retention or loss of written complexity after their study abroad trip. With the data at hand, we can assess that a gain in written complexity has been made. The degree to which it increased depends on the factors listed above. The learners produced fewer T-Units and clauses, yet those T-Units had greater clausal diversity and increased length.

RQ 4, “How will the average score ranges of writing accuracy between the first two years (Year A & Year B) differ from the final two years (Year B & Year C)?”. Results indicate that there is improvement during the latter years. The number of fully correct T-Units showed major improvement between the last two years, whereas fully-correct clauses showed less improvement within the last years. I attribute this to the grammar present in Year B, because one incorrect clause would render an entire T-Unit to be incorrect, despite having other correct clauses within. When looking at the error-level analysis, it is important to analyze which error was the most common for each year. Although more errors were produced in Year C, most of them were Level 1 errors; Year B contained more Level 2 errors. Year C had the lowest number of Level 3 errors, which steadily decreased each year of the study.

Lastly, the WCR scores showed major increases during the last two years. These scores increased by hundreds of magnitudes more in size those of the first two years. This shows the improvements made in accuracy after the study abroad period. Similar to the complexity comparisons, it would be insightful to see how learners perform in subsequent years after their study abroad. This way, skill retention or loss could be analyzed. Each of the measures, when compiled with the WCR scores, further reflect this strong increase in written accuracy in the last two years.

RQ 5, “Based on the measures in this study, how do any changes in complexity and accuracy over the three-year timespan compare to one another?”. Results of the study confirm the statements in Thewissen (2013) and Ortega (2003) that not all learning is linear. Certain performance aspects will take precedence over others at given points in time. These statements are particularly true here, as the increases found in the data were not uniform across all measures. Comparing the number of T-Units and clauses to the number of fully-correct T-Units and clauses shows opposing changes between them. The average number of T-Units and clauses decreases as the number of fully-correct T-Units and clauses increases during the study. The remaining measures of complexity indicate increases, although not to the same degree as the accuracy measures. With this data, we can conclude that written accuracy improved more than written complexity, although both indicated signs of improvement.

## 5.2 Study Abroad and Language Development

Studies that have reported the beneficial effects of study abroad programs often did so with spoken data from students. However, as mentioned in DeKeyser (2007), there were very few studies pertaining to the impact of a study abroad on the written proficiency of a learner. Although this study concerned itself with the development of written accuracy and complexity over three years, it is still insightful to the impact study abroad has on written language development. Interestingly, these results contradict the findings of Pérez-Vidal & Juan-Garau (2009), where the written fluency and complexity increased, but the accuracy of the writing samples decreased. Although the current study did not work with written fluency, it did show major improvements in accuracy and moderate improvements in complexity.

Methodological differences are important to note here, as the accuracy measure used in Pérez-Vidal & Juan-Garau (2009) was errors per word in the text, and the complexity was a split between both lexical and grammatical complexity. The current study used WCR, which is inherently a measure of communicative accuracy, and the complexity measures, both internal and supplementary, relate to the textual complexity of the samples. Had the accuracy measure used for this study been errors per word in the text, then the results would have been different. This is due to clauses, at times, containing more than one error, which WCR does not account for. Additionally, if lexical complexity were to be analyzed, I would hypothesize that a larger effect on complexity would be visible, as many samples in the years of the study appeared to contain more diverse lexical items than the previous years. Returning to this study later with a different set of accuracy and complexity measures may yield interesting results.

Regarding the language development of learners over the three years, the written complexity and accuracy improved, yet not directly in line with one another. One aspect that was not covered in this study was the potential impact of a language-learning plateau. As a reminder, the language-learning plateau, which was often seen in EFL learners, is where language progression of a learner significantly decreases as compared to the earlier stages of language acquisition. This is seen between the intermediate and advanced levels, as mentioned in Richards (2008) and Shormani (2013). This decrease in language acquisition has been attributed to many factors, such as rules in the language already being acquired by learners at earlier stages of learning, or learner motivation in the language decreasing as time goes on, as mentioned in Mirzaei et al. (2017). Due to the samples of the current study spanning three years, the likelihood of a language-learning plateau exhibited by the results here is possible, as the learners are expected to progress from a lower-intermediate proficiency level to a lower-advanced level by the third year. However,

I believe that, based on the data in the study, it will be beneficial to determine if either complexity or accuracy have any plateaus, and where they occur.

The complexity data displayed large increases initially during the intermediate years but showed only minor changes in complexity during the later years. These years were when the learners were at an advanced proficiency level (C1/C2). The smaller magnitudes of changes seen during the later years lends support to the idea of a plateau occurring in written complexity after the study abroad. An important thing to note here is that the current study did not concern itself with the changes in lexical complexity. Had lexical complexity been a measure of analysis, then the location of the plateau may vary. The need to divide syntactic and lexical complexity would be more apparent, as the two could experience differing areas of plateauing.

In stark contrast, when referring to the distribution of the error levels and WCR scores across the three years, a plateau is found between the first two years, not the final two years. Due to the plateau occurring before the study abroad, this lends evidence to the benefits of study abroad, as it allowed for learners to continue progressing beyond the plateau that they had been at for two years. The occurrences of the plateaus for complexity and accuracy pose interesting results, as learners conducting a study abroad would most likely gain more motivation in their language studies, and they should have reduced difficulty producing the language after being abroad for six to twelve months. This would result in a plateau towards the beginning of the study for both complexity and accuracy.

To return to Thewissen (2013) and Ortega (2003), learning is not linear, and certain aspects of performance take precedence over another. When considering these statements, it is possible that the study abroad was more beneficial for the development of accuracy than complexity, due to the timing of the plateaus. However, not all complexity measures used in this study showed

decreases after the study abroad, such as the average length of T-Units. To further analyze the possible effects of a language plateau, and where they occur, a study concerning the longitudinal effects of a study abroad on complexity and accuracy could be beneficial.

### 5.3 Resource Management in Writing

Referring to Harrington & Sawyer (1992) with Working Memory Capacity (WMC), the larger and more complex the task is that learners are required to complete, the more demanding the task will be on the mental resources of the learner. This is reflected in the data of this study when analyzing the trade-offs between accuracy and complexity in the samples over the three years. This trade-off, as reflected in Skehans' and Robinsons' assessments of working memory (2015, 2010), is seen in these tasks: learners often increase in complexity as they decrease in accuracy, but they decrease in accuracy when they improve in complexity.

Regarding task types, the tasks were broken up into three categories: *Aufsatz* (General, argumentative essay), *Kritische Zusammenfassung* (Critical Summary), and *Kritische Kommentar* (Critical Commentary), with tasks in order of complexity. During Year A, learners primarily produced samples in the area of *Aufsatz*, and then *Kritische Zusammenfassung* in Year B, leaving the most complex task of *Kritische Kommentar* for Year C. When comparing the different levels of task complexity to the progressions of complexity and accuracy, it appears that the increase in task complexity from a general sentence did not negatively impact written complexity, since the complexity of samples in Year B showed improvement over those in Year A. Although the learners were required to refer to a text while writing their samples during the second year, I believe that the demands on learner resources in this task were much less than the general essays in Year

A. This is due to learners being able to refer to the text for help with structures. However, samples from Year A consisted solely of learner knowledge, without supplementary texts.

The accuracy showed minor changes during the first two years. I attribute this to the possibility of learners using unfamiliar grammar found in the required texts. This resulted in more severe errors that negatively impacted accuracy. An example of this was the large number of errors present when students tried to use the "zu- *Infinitive*". Interestingly, the accuracy of the learners improved with the more complex tasks found in Year C. They were provided a supplementary text to help during the writing process, yet it was not required for learners to refer to this text while they wrote. I believe that, due to having access to the text, learners once again used grammar presented in the text. Yet due to the increases in language proficiency, they were able to correctly use the grammar. This resulted in increased accuracy scores. It would be interesting to see how the written performances of the students after studying abroad would vary if they were to instead write additional argumentative essays without text supplementation.

The supplemental materials not only provided learners with additional grammatical structures to refer to, but they also gave them additional opportunities to consult an internal monitor to decide to continue writing. Consulting the monitor, either while reading or working with the language during the sample, refers to the concept of written languaging, in which learners reflect on their own language use, as mentioned in Ishikawa (2018). As language skills increase, the expected use of the monitor should decrease, yet may still be present. This implies that learners in this study show reduced usage of the monitor over time. However, without the use of post-test surveys asking learners about their time spent reflecting on the language, the assumptions above are mere speculation. But, as seen in Ishikawa (2018) and Sasaki (2007), the more a learner progresses in the language, the less the learner will rely on the use of an internal monitor, due to

additional confidence in the language. Employing post-test surveys asking students how they reflected on their language use would be beneficial in shedding light on this issue.

Based on the complexity of the tasks, conclusions can be drawn that supplementary materials decreased the impact on WMC. The materials would provide learners with the opportunity to use structures they normally would not, resulting in the increased complexity seen in Year B. Similarly, minor improvement to accuracy was likely due to unfamiliarity regarding the grammar. The increase in language proficiency during the study also reduces the impact of task complexity on WMC. This is because learners are more familiar with complex structures. However, it is important to note that these are assumptions, as the current study did not implore measures that specifically analyze the impact of task complexity on WMC. Additional research to confirm these conclusions is necessary.

In this chapter, the research questions were identified once more, and answered based on the data of the study. The implications of study abroad and the language-learning plateau were also discussed, with highlighted potential for future study. Lastly, task complexity was discussed, and how it impacted the Working Memory Capacity of the learners throughout the study.

## CHAPTER 6. CONCLUSION

This chapter begins with a brief summary of the results. The limitations of the study regarding the implications to study abroad methodological issues are discussed. The chapter concludes with considerations for future research, as well as uses for the data present in the study.

I began this paper addressing how learners that participate in study abroad are a population of interest in SLA. This is due to the potential for major language change during the time abroad, although the improvement in written contexts is lightly researched. As the research in this paper shows, the written complexity and accuracy of learners increased over this longitudinal period. Written accuracy particularly showed major improvements during the last two years, but the influence of study abroad on this change is unknown. It was also found that complexity and accuracy both increased, with accuracy increasing more than complexity. These results indicate improvements in written language development over the three years.

### 6.1 Limitations to the Current Study

Although this study provides useful insight into the development of accuracy and complexity over the course of three years, there are several limitations. One limitation is regarding the measures of complexity used, as it does not take lexical complexity into account. This study focused solely on syntactic complexity. While the text complexity does show the complexity of T-Units that learners produce, it does not show the development of the lexicon. This is especially relevant over the course of three years, as learners would be expected to acquire significant amounts of lexical knowledge during that time. Of course, this is also reliant on the learner's interaction with the language while abroad. Learners with minimal interactions with the language

would not show major improvements in lexical complexity. However, the best scenario for analyzing complexity would be for a study to explore both lexical and syntactic complexity, alongside an accuracy measurement.

Another limitation is the lack of a control group within the study. The development in the study abroad group's written language cannot be attributed definitively to the study abroad context. Although studies have shown study abroad to be beneficial to learner's written proficiencies, the same conclusions may not be drawn here, due to the lack of a control group. The writing proficiencies of the learners did improve, although the impact of the study abroad is unknown. This ambiguity is also present from the lack of student language samples while they were abroad. If these samples were provided, alongside a control group, then the impact of study abroad could be further identified.

The location of learners within their proficiency score ranges was also ambiguous. Although CLEG identified learners in Year A as having B1/B2 proficiency, in Year B having B2, and Year C having C1/C2, the placement of learners within these ranges was not completely defined. Some learners could have been low in B1 proficiency, whereas others could have had higher B2 proficiency. The same applies for the "post-A level" test that all learners were required to take before starting their language program. This information would have been beneficial in determining individual development over the course of the study.

Lastly, due to the samples present in the study, it was difficult to analyze the impact that task complexity had on WMC. This was due to a lack of post-test surveys or questionnaires used to gather accurate data on this issue. Although inferences on how WMC was impacted through the results and conditions of the data, this is merely speculation. This is a limitation to using corpus data, as the amount of supplementary data, such as post-test surveys, can widely vary. For a future

study, a live population will serve to flesh out these additional areas of measurement. Different forms of written feedback could further shed light on the written language learning process (Fazilatfar et. al., 2014). In a similar manner, having a live population to gather data from would allow for learner stress data to be recorded, in order to determine how tasks further affected the learner's performance through the additional taxation on WMC.

## 6.2 Considerations for Future Study & Uses

There are many directions for future research on these issues. Based on the limitations, different measures could be chosen to analyze the improvement of written complexity and accuracy. A different time frame could also be chosen as well. For example, it may be interesting to analyze the same measures, yet over a longer period, or to examine the development in these language skills long after the learner has completed a study abroad. This would analyze how well the immediate effects of study abroad are retained. Important aspects of the writing process, such as written feedback on writing samples, also contribute to how forms and rules are gathered. This is not covered in the current study, due to the lacking information regarding language samples while abroad. Analyzing the written language learning processes before and after their study abroad experiences would also yield interesting data regarding reflection of language use.

One measurement that was not looked during this study was writing fluency. Writing fluency can be very insightful to evaluating the growth in language proficiency and confidence over time. For writing fluency, measures such as the amount of time learners spend pausing, or how often learners return to correct their own work, can be used for assessment. All the mentioned avenues for study would be extremely beneficial to understanding how learners progress with their written language skills.

While there are many different directions for future research, one important aspect to mention is the use of this data outside of research purposes, namely in the field of pedagogy. As language instructors, these data are invaluable when used alongside oral data to encourage learners to study abroad. It is often the argument that a study abroad will help a learners' language skills. However, only oral production and listening comprehension are typically mentioned. With the small amount of data on increases in writing proficiency, relative to oral data, this study serves to support the claims made by other studies that writing proficiency shows signs of increases, both in written complexity and accuracy.

## REFERENCES

- CLEG Learner Corpus (2013). Lancaster University
- Allen, H. W. (2002). *Does study abroad make a difference? An investigation of linguistic and motivational outcomes* (Doctoral dissertation). Available at: <http://works.bepress.com/heatherwillisallen/50/>
- Bergman P., Abrahamsson, T. (2004). Bedömning av språkfärdigheten hos andraspråkselever. In K. Hyltenstam & I. Lindberg (Eds.), *Svenska som andraspråk – i forskning, undervisning och samhälle* (pp. 597-626). Lund: Studentlitteratur.
- Bergsleithner, J. (2010). Working memory capacity and L2 writing performance. *Ciências & Cognição, 15*(2), 2-20.
- Biber, D., Gray, B., Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*, 5-35.
- Brezina, V., Pallotti, G. (2016). Morphological complexity in L2 texts. *Second Language Research, 35*(1), 99-119.
- Bulté, B., Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken & V. Ineke (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy, and Fluency in SLA* (pp. 21-46). Amsterdam, Netherlands: John Benjamins Publishing.
- Carlson, J. S., Barber, E. G., & Burn, B. B. (1991). Study abroad: The experience of American undergraduates in Western Europe and the United States. *Center for International Educational Exchange*. ERIC No. ED 340322.
- Chambers, F. (1997). What do we mean by fluency? *System, 25*(4), 535-544.

- Crookes, G. (1990). The utterance, and other basic units for second language discourse analysis. *Applied Linguistics*, 11, 183-199.
- Crossley, S., Salsbury, T. & McNamara, D. (2009). Measuring L2 lexical growth using hypernymic relationships. *Language Learning*, 59(2), 307-334.
- Crossley, S., McNamara, D. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, 26, 66-79.
- Davie, J. (1996). Language skills, course development and the year abroad. *Language Learning Journal*, 13, 73–76.
- DeKeyser, R. M. (2007). Study abroad as foreign language practice. In R. M. DeKeyser (Ed.), *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 208-226). Cambridge, United Kingdom: Cambridge University Press.
- Ellis R. & G. Barkhuizen. (2005). *Analysing Learner Language*. Oxford: Oxford University Press.
- Evans, N., Hartshorn, K., Cox, T., & Martin de Jel, T. (2014), Measuring written linguistic accuracy with weighted clause ratios: A question of validity. *Journal of Second Language Writing*, 24, 33-50.
- Fazilatfar, A. M., Fallah, N., Hamavandi, M., & Rostamian, M. (2014). The effect of unfocused written corrective feedback on syntactic and lexical complexity of L2 writing. *Procedia – Social and Behavioral Sciences*, 98, 482-488.
- Foster, P. (2016). *Comparing classroom and laboratory task performance: An empirical study*. Manuscript in preparation.
- Foster, P., Skehan, S. (1996). The influence of planning and task type on second language performance. *SSLA*, 18, 299-323.

- Foster, P., Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics*, 36, 98-116.
- Frear, M., Bitchener, J. (2015). The effects of cognitive task complexity on writing complexity. *Journal of Second Language Writing*, 30, 45-57.
- Freed, B. F. (Ed.). (1995). *Second language acquisition in a study abroad context*. Amsterdam, Netherlands: John Benjamins Publishing.
- Freed, B. F., Segalowitz, N. & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26, 275–301.
- Freed, B. F., So, S., & Lazar, N. A. (2003). Language learning abroad: how do gains in written fluency compare with oral fluency in French as a second language? *ADFL Bulletin*, 34(3), 34-40.
- Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics*, 45, 215-240.
- Gilabert, R., Baron, J., & Llanes, M. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *IRAL*, 47, 367-395. <https://doi.org/10.1515/iral.2009.016>
- Golonka, E. M. (2001). Identification of salient linguistic and metalinguistic variables in the prediction of oral proficiency gain at the Advanced-Level Threshold among adult learners of Russian. Ph.D. Thesis, Bryn Mawr College, *Dissertation Abstracts International*, 61.
- Harrington, M., Sawyer, M. (1992). L2 reading capacity and L2 reading skill. *SSLA*, 14, 25-38.
- Hammerly, H. (1990). *Fluency and accuracy: Toward balance in language teaching and learning*. Clevedon: Multilingual Matters.

- Hawkins, R. (2001). *Second language syntax: A generative introduction*. Oxford, UK: Wiley-Blackwell.
- Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. Chicago, Illinois: National Council of Teachers of English.
- Ishikawa, T. (2006). The effect of task complexity and language proficiency on task-based language performance. *The Journal of Asia Tefl.*, 3(4), 193-225.
- Ishikawa, M. (2018). Written languaging, learners' proficiency levels and L2 grammar learning. *System*, 74, 50-61.
- Iwashita, N. (2006). Syntactic complexity measures and their relation to oral proficiency in Japanese as a foreign language. *Language Assessment Quarterly*, 3(2), 151-169.
- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, 24(4), 541-577.
- Kitiabi, D. B. (2018). *Second Language Acquisition in a Study Abroad Context: International Students' Perspectives of the Evolution of their Second Language Self'* (Unpublished doctoral dissertation). Bowling Green State University, Bowling Green, OH.
- Kozlova, I., Presas, M. (2016). Resolving L2 written text production problems using reference sources. In L. Molina & L. Santamaria (Eds.), *Traducción, Interpretación y Estudios interculturales*, Granada: Editorial Comares, Colección Interlingua.
- Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. In S. Foster-Cohen, M. P. Garcia-Mayo & J. Cenoz (Eds.), *Eurosla Yearbook* (Vol. 5, pp. 195-222). Amsterdam, Netherlands: John Benjamins Publishing.

- Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching*, 45(3), 261–284.
- Kuiken, F., & Vedder, I. (2007). Task complexity, task characteristics and measures of linguistic performance. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard & I. Vedder (Eds.), *Complexity, accuracy and fluency in second language use, learning and teaching* (pp. 113–126). Brussels, Belgium: KVAB.
- Lee, Y. G. (2002). Effects on task complexity on the complexity and accuracy of oral production in L2 Korean (Unpublished doctoral dissertation). University of Hawaii at Manoa, Honolulu.
- Leonard, K. R., & Shea, C. (2017). L2 speaking development during study abroad: Fluency, accuracy, complexity, and underlying cognitive factors. *The Modern Language Journal*, 101(1), 179-193.
- Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.
- Lu, Y. (2010). Cognitive factors contributing to Chinese EFL learners' L2 writing performance in timed essay writing (Unpublished doctoral dissertation). Georgia State University, Georgia.
- MacIntyre, P., Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44(2), 283-305.
- Mirzaei, M., Zoghi, M., & Asl, H. D. (2017). Understanding the language learning plateau: A grounded-theory study. *Teaching English Language*, 11(2), 195-222.

- Nagy, W. (2005). Why vocabulary instruction needs to be long-term and comprehensive. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and Learning Vocabulary: Bringing Research to Practice* (pp. 27–44). Mahwah, NJ: Lawrence Erlbaum.
- Nitta, R., & Baba, K. (2014). A longitudinal study from a dynamic systems perspective. *Task-Based Language Learning – Insights from and for L2 Writing*, 7, 107-136.
- Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555-578.
- Nuevo, A. M. (2006). Task complexity and interaction (Unpublished doctoral dissertation). Georgetown University, U.S.A.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college- level L2 writing. *Applied Linguistics*, 24(4), 492-518.
- Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590-601.
- Pérez-Vidal, C., & Juan-Garau, M. (2009). The effect of study abroad (SA) on written performance. *EUROSLA Yearbook*, 9(1), 269-295.
- Polat, B., & Kim, Y. (2013). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics*, 35(2), 184-207.
- Polio, C., & Shea, M. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10-27.
- Révész, A., Michel, M., & Gilabert, R. (2016). Measuring cognitive task demands using dual-task methodology, subjective self-ratings, and expert judgments. *Studies in Second Language Acquisition*, 38, 703–737.

- Richards, J. C. (2008). *Moving beyond the plateau: From intermediate to advanced levels in language learning*. New York: Cambridge University Press.
- Riggenbach, H. (1991). Toward an understanding of fluency: a microanalysis of nonnative speaker conversations. *Discourse Processes, 14*, 423-441.
- Robinson, P. (1995). Attention, memory, and the “noticing” hypothesis. *Language Learning, 45*(2), 283-331.
- Sasaki, M. (2007). Effects of study- abroad experiences on EFL writers: A multiple- data analysis. *The Modern Language Journal, 91*(4), 602-620.
- Shormani, M. Q. (2013). Fossilization and plateau effect in second language acquisition. *Language in India, 13*(2), 763-784.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.
- Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360-407.
- Storch, N., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M. P. García Mayo (Ed.), *Investigating Tasks in Formal Language Learning*. Bristol, UK: St. Nicholas House.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253), Rowley, MA: Newbury House.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 97-114). Oxford, UK: Oxford University Press.

- Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus. *The Modern Language Journal*, 97(1), 77-101.
- Tsiriotakis, I., Vassilaki, E., Spantidakis, I., & Stavrou, N. (2016). The examination of the effects of writing strategy-based procedural facilitative environments on students' English foreign language writing anxiety levels. *Front Psychol*, 7, doi: 10.3389/fpsyg.2016.02074
- VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12(3), 287-301.
- VanPatten, B., & Oikkenon, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition*, 18, 495-510.
- Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67(2), 394-430.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity*. Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center
- Wright, C. (2018). Research in memory and processing in second language acquisition (pp. 203-220). In Wright, C., Piske, T. & Martha, Y. (Eds.), *Mind Matters in SLA*.
- Yager, K. (1998). Learning Spanish in Mexico: The effect of informal contact and student attitudes on language gain. *Hispania*, 81(4), 898-914.
- Yi, B., & Luo, S. (2013). Working memory and lexical knowledge in L2 argumentative writing. *Asian Journal of English Language Teaching*, 23, 83-102.
- Zheng, Y. (2008). Anxiety and second/foreign language learning revisited. *Canadian Journal for New Scholars in Education*, 1, 1-12.

## APPENDIX A: TASK TYPE EXAMPLES

### *Aufsatz* Example (Learner ID: B1003\_05)

Seit vielen Jahren hat diese Argument weiter gegangen; für die Mehrheit der arbeitenden Bevölkerung es scheint, dass wir um zu leben arbeiten. Diese Ansicht vorherrscht wegen der langen Arbeitszeit, man muss am Wochenende arbeiten und dafür haben wir sehr wenige freie Tage. Aber wenn wir eine Wahl haben würden, würde die gegenwärtige Lage bleibt die gleiche? Die wahrscheinliche Antwort ist "Nein", natürlich mögen wir mehr Freizeit, aber ist es machbar? Die Franzosen vertreten für die 35 Stunden Woche, sie bestehen, dass es keine Notwendigkeit gibt zu arbeiten mehr. In Wirklichkeit würde die Wirtschaft leiden, weil das Land weniger produktiv sein würde. Natürlich wenn wir weniger arbeiten, würden die Stressniveaus nehmen ab. Auf der einer Seite würden wir arbeiten wenige Stunden aber auf der anderen Seite würden wir wenige Löhne bekommen und zweifellos würde der Lebensstandard verschlechtern, jedoch würde die Qualität des Lebens verbessern. Das heisst: wir würden glücklicher werden wegen der kurzen Arbeitszeit selbst wenn wir wenige Geld bekommen würden. Es gibt gewissen Leute, die sagen, dass wenn sie die Lotterie gewinnen würden, würden sie noch arbeiten, aber ist es richtig? Ich glaube, dass wenn ich die Lotterie gewonnen hatte, würde ich nicht arbeiten, weil es mehr zum Leben als arbeiten und Geld würde keine Rolle spielen! Ein positiver Aspekt von kurzen Arbeitszeiten ist, dass es weniger Arbeitslosigkeit würde sein, weil Leute weniger Stunden arbeiten also gibt es mehr Arbeitgelegenheiten. Es wird allgemein angenommen, dass Leben mehr als Arbeit umfasst. Die Briten brauchen mehr Freizeit um mehr Zeit mit ihren Familien zu verbringen, Hobbys machen und für Entspannung. Um einer glücklicher Arbeiterschaft zu schaffen, muss die Regierung Freizeit betont lieber als mehr Stunden bei der Arbeit verbringen. Wir arbeiten nur das Geld machen, also könne wir machen was wir wollen d.h im Urlaub fahren, ins Restaurants gehen u.s.w. Durch Arbeit man fühlt sich wertvoll, man bekommt Zufriedenheit am Arbeitsplatz und ein Leistungsgefühl. Arbeit gibt man ein Lebensziel und oft ist eine Stelle ein Statussymbol. Die ideal Lage ist ein Gleichgewicht zwischen Arbeit und Freizeit, leider ist diese Gleichgewicht mehr einseitig, mit Arbeit dominieren unseres Leben.

*Kritische Kommentar* Example (Learner ID C1008\_04)

Sollten gewalttätige Computerspiele für Jugendliche verboten werden? Gewalttätige Computerspiele sind schon seit Jahren ein umstrittenes Thema. Es gibt Menschen, die die gerne verbieten würden, obwohl sie sehr populär unter Kindern sind. Der vorliegende Text handelt sich um die Meinungen der Eltern, bzw. Psychologen und Wissenschaftler, und fragt wer ist dafür verantwortlich, wenn gewalttätige Computer eine negative Wirkung haben, und ob das stimmt überhaupt. Es ist natürlich, dass Eltern mit Sorge erfüllt werden, wenn ihre Kinder von gewalttätige Computerspiele fasziniert sind. Aber wenn Psychologen sagen, dass sie Kindern schaden, kann man diese Meinung für noch ernster halten. Was Kinder sehen, möchten sie oft imitieren, deshalb wird es schwer für Eltern, deren Kinder zu kontrollieren, wenn es gewalttätige Computerspiele zum Kaufen gibt. Es ist auch schwer für Kinder zu verstehen, dass Computerspiele schädlich seien, wenn sie nicht verboten sind. Trotzdem könnte man sagen, dass Eltern zu wenig verantwortlich sind. Wenn die Kinder Computerspiele kaufen oder bekommen, dürfen Eltern entscheiden, ob sie damit spielen. Laut Statistik im Text verbringen Eltern wenig Zeit, Computergames mit ihren Kindern zu spielen. Wenn sie besser wissen würden, genau was ihre Kinder machen und was die Spiele beinhalten, könnten sie besser erklären, warum sie schädlich sind. Sonst könnten sie ihre Kinder ermutigen, andere Hobbys zu haben, die man als Familie macht. Vielleicht wäre es eine sehr gute Idee, wenn Familien noch mehr Zeit verbringen würden. Ein Verbot von gewalttätige Computerspiele könnte es ermöglichen, entweder dass Familien mehr Zeit zusammen haben, oder dass Kinder Computerspiele kaufen, die dem Lernen helfen. Wenn Eltern freiwillig strenger wären, wäre ein Verbot von gewalttätigen Computerspielen sehr effektiv. Auf der anderen Seite wissen wir nicht, dass Gewalt in Computerspielen immer negative Folgen hat. Es gibt immer noch zu wenige Studien, die das Thema untersucht haben. Auch wenn eine Mehrheit von Experten einig sind, dass es Kindern schädlich ist, brutale und blutige Bilder zu sehen, stimmen sie zu, dass es nicht nur Bilder sind, die Aggressivität in Kindern verursachen. Jedenfalls sehen sie solche Bilder im Fernsehen und im Kino, nicht nur in den Computerspielen. Obwohl sie nicht die einzige Quelle von Gewalt sind, Computerspiele können zu Gewalttätigkeit in Menschen führen, wenn sie oft solche Spiele treiben. Ein Beispiel davon ist der Fall von Columbine High School, in welchem die Täter eines Blutbades wollten, dass ihrer Anschlag das Spiel Ego-Shooters imitieren würde. Wenn Kinder beginnen, einander zu töten, weil sie daran gewöhnt sind, das oft und strategisch zu machen,

muss man ein Verbot in Betracht ziehen. Es gibt noch viel mehr Argumente dafür und dagegen, dass Computerspiele für Jugendliche verboten werden. Es stimmt, dass Eltern eine wichtige Rolle spielen und sie besser kontrollieren, wie ihre Kinder die Freizeit verwendet, aber es wäre unmöglich und unhilfreich, das die ganz Zeit zu machen. Obwohl die Ergebnisse der wenigen Studien nicht eindeutig sind, weiss man schon, dass Blut und Gewalt im Fernsehen negativer Folgen für Kinder haben. Eltern und die Hersteller von Computerspielen sind gleichzeitig schuld, wenn Schulkinder erschossen werden. Wir sollten alle verantwortlicher sein, deswegen brauchen wir ein Verbot von gewalttätigen Computerspielen.

*Kritische Zusammenfassung Example (Learner ID C1005\_07)*

*Das Plagiat bei den Universitäten* Der Artikel handelt sich um das Problem von Plagiat bei den irischen Universitäten. Akademiker aus den Universitäten werden miteinander treffen um zu diskutieren, wie das Problem zu konfrontieren, die derzeitigen Regeln nachzuschauen, ein neues System zu erwägen und ihre weltweiten Reputationen zu schützen. Zwei selbständige Studien berichten, dass mehrere Studenten haben in Kopieren mit dem Internet als Hauptquelle aufgenommen. Einigen Akademikern stimmen die Ergebnisse nicht zu. Im Moment haben die Institutionen ihre eigene Plagiatpolitik aber sie wollen es weiter nehmen. Es gibt verschiedene mögliche Bestrafungen bis zum Ausschluss für Studenten die plagieren aber es kann schwierig sein um ein Fall zu beweisen. Die Studenten müssen erklären die Aufgaben sind ihre eigene Arbeit und Quellen bemerken. Einige Universitäten benutzen spezialisierte Computerprogramme um kopierte Text in Studentenaufgaben zu finden, Geänderte Stil oder Arbeitsqualität kann auch Plagiat kennzeichnen. Laut dem Präsident eines Studentenwerks, wissen nicht alle die Studenten die Regeln über Plagiat. Ein Akademiker aus Oxford beschuldigt die Schule.

**APPENDIX B: DATA GRAPHS**







