ELEMENTARY ELL INTERACTION: MAINSTREAM V. SHELTERED INSTRUCTIONAL SETTINGS

By

JOAN ANN JOHNSTON

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Teaching and Learning

MAY 2013

© Copyright by Joan Ann Johnston, 2013
All Rights Reserved
To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of Joan Ann Johnston find it satisfactory and recommend that it be accepted.

_________________________________________
Thomas Salsbury, Ph.D., Chair

_________________________________________
Brian French, Ph.D.

_________________________________________
Mary Martha Savage, Ph.D.
ACKNOWLEDGEMENTS

I would like to thank the following individuals who assisted in the collection of data by documenting and video taping observations in language camp classrooms: Amy Berube, Sawadipong Bavornvanijyakul, Liliana Restrepo, and Ming Tsai. I would especially like to thank Frances Manring and Dani Troupe for, not only assisting with documenting mainstream observations, but for additionally helping to pilot test the observation sheets, troubleshoot the video equipment, and assist with the reviewing and triangulation of data from the videos. All six of the assistant observers were of great assistance in this process.

I would also like to thank the members of my dissertation committee, Thomas Salsbury, Brian French, and Martha Savage, for providing the support and feedback to complete this study. Their input throughout the process has been invaluable.

Finally, I would like to acknowledge the classroom teachers and English language development specialists who allowed me into their classrooms to complete this study. Their openness in sharing their classrooms and their thoughts are greatly appreciated.
ELEMENTARY ELL INTERACTION: MAINSTREAM V. SHELTERED INSTRUCTIONAL SETTINGS

Abstract

by Joan Ann Johnston, Ph.D.
Washington State University
May 2013

Chair: Thomas Salsbury

This mixed methods study investigated the English verbal interactions of seven 3rd-6th grade beginning level English language learners across three different instructional settings: the mainstream grade level classroom, the sheltered English classroom, and the ELL pullout group. The quantitative component of the study documented significant differences in the frequency of English verbal interactions between the three instructional settings. The students were found to have the lowest frequency of overall English interaction in the mainstream classrooms, with the highest rates of interaction occurring in the ELL pullout groups. Peer interactions were found to be highest in the sheltered English classrooms, where students were among English language learning peers in a 3-week summer language program. Peer interactions in the mainstream classrooms, where the majority of students were native English speakers, were significantly lower than either the sheltered classroom or the pullout group.

Grounded analysis of the observation notes from each observation session revealed completely divergent themes between the mainstream and sheltered classrooms. Students in the sheltered classrooms were observed to be highly engaged and enthusiastic; whereas, the same students were found to be largely disengaged, isolated, and frustrated or angry in the mainstream classrooms.
Taken together, the quantitative and qualitative findings indicate that beginning level students are likely to be engaged in the content and the language when surrounded by other English language learners in the sheltered classrooms. Conversely, there is a strong tendency for students to withdraw from interaction or shut down when competing academically and socially with native speakers of English in the mainstream classroom.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iv-v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>x</td>
</tr>
<tr>
<td><strong>CHAPTERS</strong></td>
<td></td>
</tr>
<tr>
<td>1. AN INTRODUCTION TO THE STUDY</td>
<td>1</td>
</tr>
<tr>
<td>2. LANGUAGE DEVELOPMENT, INSTRUCTIONAL MODELS, AND MIXED METHODS RESEARCH</td>
<td>7</td>
</tr>
<tr>
<td>A Brief History of English Immersion</td>
<td>8</td>
</tr>
<tr>
<td>Theories of Second Language Acquisition</td>
<td>13</td>
</tr>
<tr>
<td>Sheltering and Instructional Strategies</td>
<td>28</td>
</tr>
<tr>
<td>Into The K-12 Classroom: What the Research Says</td>
<td>47</td>
</tr>
<tr>
<td>History of Mixed Methods Design</td>
<td>53</td>
</tr>
<tr>
<td>The Current Study</td>
<td>67</td>
</tr>
<tr>
<td>3. OVERVIEW OF METHODS</td>
<td>69</td>
</tr>
<tr>
<td>4. QUANTIFYING INTERACTION FREQUENCY</td>
<td>80</td>
</tr>
<tr>
<td>Analysis and Results</td>
<td>82</td>
</tr>
<tr>
<td>5. WHAT’S BEHIND THE NUMBERS? A QUALITATIVE ANALYSIS OF CLASSROOM OBSERVATIONS</td>
<td>97</td>
</tr>
<tr>
<td>Methods</td>
<td>98</td>
</tr>
</tbody>
</table>
Divergent Themes in the Findings .......................................................... 102
Qualitatively Opposed: A Summary ....................................................... 119

6. DISCUSSION AND CONCLUSIONS .................................................. 122

BIBLIOGRAPHY ......................................................................................... 134

APPENDICES

A. STRATEGIES FOUND IN CALLA, SIOP, AND GLAD ....................... 147
B. RESEARCH OBSERVER EQUIPMENT/MATERIALS PROTOCOL ........... 148
C. OVERVIEW OF DATA COLLECTION .................................................. 149
LIST OF TABLES

1. Comparison of Previous to Current TBIP Definition .................................................. 31
2. All Participants by Instructional Setting, N=2 ......................................................... 75
3. Multivariate Tests ........................................................................................................ 84
4. Dependent (Paired Samples) T-test ............................................................................. 85
5. Dependent (Paired Samples) T-test for Teacher-Student Interaction Rates ................. 87
6. Dependant T-test for Peer (Student-Student) Interaction Rates Between Settings ......... 90
7. Paired Samples Test for Teacher-Student v. Student-Student Interactions Within Settings....92
8. Individual Students ...................................................................................................... 99
9. Overview of Data Collection ........................................................................................ 149
LIST OF FIGURES

1. Legislation and Court Cases Affecting the Education of Language Minority Students……12
2. Categorization of Communication Strategies as Outlined by Tarone.........................21
3. The Negotiation Process..........................................................................................23
4. Input, comprehension, output process......................................................................25
5. Participants by Observational Settings, $N=20$.........................................................74
6. Sequence of Observations.........................................................................................78
7. Mean Interaction Rates by Instructional Setting.........................................................86
8. Teacher-Student Interactions in Mainstream and Sheltered Settings.........................88
9. Comparison of Peer Interaction Rates by Setting.....................................................91
10. Teacher and Peer Interactions as a Percent of Overall Interactions..........................93
11. Individual Student Interaction Rates.......................................................................95
12. Categories of Coded Data from Mainstream Observation Notes.............................103
13. Wordle of Dominant Words in the Mainstream Observation Notes.........................107
14. Positive and Negative Categorization of Coded Observation Notes.........................120
15. Wordle of Mainstream Coded Notes Excluding "Assessment".................................121
16. Wordle of Sheltered Coded Notes...........................................................................121
17. Wordle of Strategies Found in CALLA, SIOP, and GLAD......................................147
DEDICATION

This dissertation is dedicated to my husband, Dave, and my daughter, Jenny,
who have provided encouragement and support throughout this journey.

And to my students, from whom I continue to learn and grow.
Chapter 1

An Introduction to the Study

Having been a teacher of English language learners (ELLs) since 1986, I have had the opportunity to observe students of varying ages and at varying stages of language acquisition in different instructional environments. For 11 of the past 14 years I have also had the good fortune to be involved as both a teacher and an administrator in a 3-week summer language camp. In my position as an English language development (ELD) specialist in the public school setting, it was often the case that I would see the same students in the camp that I saw in the mainstream classrooms, with one big difference. I noticed that the students seemed to be much more involved when they were among their ELL peers during the language camp than when they were in their mainstream classrooms. This was likewise something that I had noticed when I taught in a sheltered program in the middle school setting, in which I recall a science teacher commenting what a quiet student Yuriy was, when, in my sheltered ELL classroom, he never stopped talking. (I had to clarify if we were, in fact, talking about the same student.) These observations made me wonder: Are there actual differences in student interaction between the two settings? If so, what other differences exist between these settings that might be influencing the different student behaviors?

Background

The number of English language learners (ELLs) being served in the state of Washington has more than tripled in the past 20 years, increasing 241% between 1990 and 2010. In that same period, the overall enrollment in Washington public schools only increased by 23%. According to the state’s Transitional Bilingual Instructional Program (TBIP), the number of students
qualified for ELL services increased by 5.7% between 2010 and 2011, nearly triple the 2% increase reported the previous year. During this same period, overall enrollment in Washington schools actually declined by 0.4% for the third consecutive year (Malagon, McCold, & Hernandez, 2011b). On the national level, just short of five million students were enrolled as ELL in U.S. public schools in the 2009-2010 school year, reflecting 10% of the total public school population for that year (U.S. Department of Education, 2010). Some estimates project that over the next twenty years 30% to 50% of students in U.S. schools will speak a language other than English in their home, many of whom will be in need of English language services upon entering school.

A Matter of Importance

The 1998 passage of California’s Proposition 227 required all public school instruction be conducted in English and limited the placement of students in sheltered English programs to one year (American Institutes for Research and WestEd, 2006). A virtually identical proposition (proposition 203) was passed in Arizona in 2000. With the passage of these two propositions, it became increasingly prevalent for even beginning level English language learners to be placed in mainstream English speaking classrooms often times from the very first day of their entry into the American school system. Typically, Washington State English language learners in the middle school and high school grades receive 1 to 4 hours per day of sheltered English instruction and in some large urban districts may even have the opportunity to attend a newcomers’ center for the first 1 or 2 semesters. In contrast, students in elementary classrooms (75% of the all eligible ELLs) typically receive little if any additional support. Districts rely instead on classroom teachers to provide much of the support for language development.
Although districts throughout the country increasingly rely on classroom teachers to serve the needs of this growing population, the teachers are provided with little training to do so. Title III of NCLB requires districts to provide professional development for teachers and staff working with ELLs in the classroom. However, the reality of this training is that it lacks in both accessibility and quantity. In 2002, the National Center for Education Statistics reported that, of the 41 percent of teachers in the U.S. who worked with English language learners in their classrooms, less than 13 percent had received any training or professional development to do so (in Malagon & Chacon, 2009, p. 6). In 2007, Crawford and Krashen (2007) reported that the average classroom teacher working with ELLs in U.S. schools had received only 4 hours of ELL training over the previous 5 years. That’s less than 50 minutes per year, equivalent to one class period or one staff meeting.

Despite increased efforts to better train mainstream teachers to work with ELLs in their classrooms, the assumption is that students will learn English more quickly if they are immersed in the language for 6 hours every day and are forced to use it. This assumption seems to be one that is shared by the general population, which Crawford (1992) labels “the immersion fetish—the idea that maximum exposure and maximum will are what counts in language acquisition” (p. 209). Harklau (1994) makes a similar observation and challenges the “folk belief that children will learn English faster if they are in regular classes with native speakers of English” (p. 242). Do students really use English more in the mainstream classroom? And, if not, why not?

There have been several studies of English language learners in the mainstream classroom at the middle school and high school levels (Duff, 2001; Harklau, 1994; Miller, 2000; Pappamihiel, 2002). Harklau (1994) found that high school English language learners were less engaged in the mainstream classrooms than in the sheltered ESL classroom, with fewer
opportunities for comprehensible input and comprehensible output. Miller (2000) had similar findings with high school immigrant students in Australia. Miller states that students had more real opportunities to use English (in an ESL reception program) than in their mainstream high schools…The irony is that moving into a mainstream high school actually limited their chances to use English. (The students) stopped trying to speak English. (p. 96)

Duff (2001) found students had “greater opportunities to interact using English” in the ESL setting than in the mainstream high school classroom, as well as exhibiting higher levels of confidence (p. 118). Pappamihiel (2002) focused primarily on issues of anxiety for 6-8 grade ELLs in the mainstream and ESL classroom, finding that students displayed higher anxiety and less interaction in the mainstream than in the ESL setting.

Although there have been numerous studies on interaction at the secondary level, particularly with 9-12 grade, there is little in the body of research regarding mainstreamed students in the k-6 grade levels. Because students in the elementary grades make up more than 75% of students qualified for ELL services in the United States, it’s important that we investigate this segment of the population and determine how best to address their language learning needs.

**Purpose**

The purpose of this comparative mixed methods study was to investigate, quantitatively, the frequency of beginning level 3rd-6th grade English language learners’ (ELLs’) verbal interactions across three different instructional settings: the mainstream grade level classroom, the ELL sheltered classroom, and the ELL pullout group. The qualitative component of the study was undertaken to provide a better understanding the contextual influences affecting ELLs’
interactions in the mainstream and sheltered classrooms. By employing both quantitative and qualitative investigative methods in a concurrent embedded design, I am able to provide evidence which contributes to answering the broader question of which instructional environment provides the best opportunity for beginning level students to interact and engage in the classroom.

**Research Questions and Data Collection**

To fulfill the purpose of this study, I posed four research questions for the quantitative component and two research questions for the qualitative component.

**Quantitative Research Questions:**
1. What are the differences in frequency of overall English verbal interactions between the mainstream, sheltered, and pullout settings?
2. How does the frequency of teacher-to-student English verbal interaction compare between the three instructional settings?
3. How does the frequency of student-to-student (peer) interactions compare between the three instructional settings?
4. Are there significant differences between teacher-to-student and student-to-student interaction frequency within each setting?

**Qualitative Research Questions:**
1. How did the students’ behavior differ between the mainstream and the sheltered instructional settings?
2. What circumstances in the classrooms may have influenced the observed interactions?

To answer these questions, data was collected using program records, structured field observations, anecdotal observation notes, and video recording. Although not systematically
analyzed, debriefing sessions, individually for teachers, and in groups, with interpreter, for students, provided opportunities for added insights into the qualitative data.
Chapter 2
Language Development, Instructional Models, and Mixed Methods Research

“Language learning evolves out of learning how to carry on conversations.”

Introduction

_Bilingual Education, Sheltered Instruction, English Immersion, and English Submersion,_ are all terms that give rise to a very heated and longstanding debate in the world of English language instruction. This is true not only in the United States but throughout the world, especially in those countries that have significant immigrant and refugee populations such as the UK, Australia, and Canada (Bureau of Population, Refugees, and Migration, 2009). In the United States, the question of whether English language learners should learn English in the mainstream classroom or in classrooms separate from the mainstream continues to be at the forefront, alongside an equally heated debate over the role and existence of bilingual education programs.

The 2001 reauthorization of the Elementary and Secondary Education Act (ESEA), in the form of No Child Left Behind (NCLB), heightened the focus of legislators, school districts, and the public on the education of English language learners in the United States. Because of increased levels of accountability currently required in federal and state legislation, more and more schools are investigating how to best meet the needs of their increasing numbers of English language learners, a demographic expected to constitute one-third of the U.S. student population by the year 2043 (Crawford & Krashen, 2007).
A Brief History of English Immersion

Issues of language and education are not as new as some would expect. They have been a part of the American educational landscape since long before the United States became a country. Crawford (1992), in his book *Hold Your Tongue*, explains the stance that Benjamin Franklin took against German language education in Pennsylvania as early as the 1750s, a stance that cost him his seat in the Pennsylvania Assembly in 1764. The education of Native American children and the issue of English language immersion goes back even further with the mission schools established along the Great Lakes in the 1600s, followed by the BIA (Bureau of Indian Affairs) boarding school era which dates back to 1879. For the past 300 years, the education of language minority children has been discussed and debated in courtrooms, congress, and the classroom.

To better understand the mainstreaming of English language learners into today’s English only classrooms, it may prove helpful to review the history of legislation and significant court cases that have affected the educational system in the United States and, in turn, have affected the education of English language learners in particular over the past 60 years. I will begin my focus with the case of Brown v. the Board of Education in 1954, when the racial integration of schools became law.

The 1954 case of Brown v. Board of Education in Kansas overruled the previous 1896 Plessy v. Ferguson decision, which permitted "separate but equal education for Negro children" (Brown v. Board of Education, 1954). The Supreme Court’s decision in May of 1954 ordered the desegregation of schools and established the now widely accepted principle of equal educational opportunity for all students. The court stated, "...where a state has undertaken to provide an opportunity for an education in its public schools, such opportunity is a right which..."
must be made available to all on equal terms” (n.p.). While this landmark case set out to provide more equitable access to education for African-American students, it had the added effect of beginning a long series of legislation and court cases for the rights of Native American students and language minority students as well. By mandating that schools begin the process of school desegregation, the push to integrate students planted the seed for the push we see today for mainstreaming language minority students.

In 1963, President John F. Kennedy stated: “Simple justice requires that public funds, to which all taxpayers of all races contribute, not be spent in any fashion which encourages, entrenches, subsidizes or results in racial discrimination” (in Title VI of the Civil Rights Act, 1964). The following year, as part of the Civil Rights Act of 1964, Title VI prohibited discrimination on the basis of race, color, and national origin in programs and activities receiving federal financial assistance. This legislation opened the door for the passage of the Elementary and Secondary Education Act (ESEA) of 1965 (20 U.S.C. 2701 et seq.), bringing focus to the needs of students in America’s public schools. It is indeed noteworthy that the 1965 act would be reauthorized as the No Child Left Behind Act of 2001, which continues to have a profound effect on education in the United States, particularly in regard to English language learners in the public school system.

The decade following the Education Act of 1965 brought about sweeping reforms in the areas of equity and language education. Title VII of the Elementary and Secondary Education Act, also known as the Bilingual Education Act of 1968, provided funding incentives for school districts interested in establishing programs to meet the language needs of large numbers of English language learners in the United States. A May 25, 1970 memorandum from the Department of Health, Education and Welfare (D/HEW) provided further interpretation beyond
the financial incentives of Title VII, clarifying that it was the “responsibility of school districts to provide equal educational opportunity to national origin minority group children deficient in English language skills” (Pottinger, 1970). A ripple effect from a focus on civil rights and bilingual education extended beyond the education of refugee and immigrant children. Four years after the passage of the Bilingual Education Act, the Indian Education Act was also passed, promoting the use and teaching of Native languages and cultures in Indian education programs throughout the country. The groundwork was now firmly set for the related debate on the education of other language minority students in the United States as well.

In a landmark 1974 class action lawsuit in San Francisco, Lau vs. Nichols, the US Supreme Court (Lau v. Nichols, 1974) reaffirmed the 1970 memorandum regarding denial of access and participation in an educational program due to inability to speak or understand English. In this case, the court made it clear that merely placing students in the regular English speaking classrooms without support to develop their English language skills, was in fact denying them equal access to education. Supreme Court Justice Douglas delivered the opinion of the Court:

> We know that those who do not understand English are certain to find their classroom experiences wholly incomprehensible and in no way meaningful. … Where inability to speak and understand the English language excludes national origin-minority group children from effective participation in the educational program offered by a school district, the district must take affirmative steps to rectify the language deficiency in order to open its instructional program to these students. (n.p.)

Justice Douglas further stated “There is no equality of treatment by providing students with the same facilities, textbooks, teachers and curriculum, for students who do not understand English
are effectively foreclosed from any meaningful education" (n.p.) setting forth that the “sink or swim” approach to language education was no longer acceptable. The Equal Educational Opportunity Act, passed later that year, enacted into federal law the principal of Lau v. Nichols that "... the failure by an educational agency to take appropriate action to overcome language barriers that impede equal participation by students in an instructional program” (20 USC Sec. 1703) constitutes denial of equal educational opportunity.

For the next 30+ years support for bilingual education has been under fire. In California, the June 3, 1998 passage of Proposition 227 virtually banned bilingual education except under certain special conditions and established a one-year "sheltered immersion" program for all ELL students (American Institutes for Research and WestEd, 2006). An almost identical law, Proposition 203, was passed in Arizona in 2000. The language introducing both laws is identical. “Young immigrant children can easily acquire full fluency in a new language, such as English, if they are heavily exposed to that language in the classroom at an early age” (California Secretary of State, 1998, section 1; Arizona Secretary of State, 2000, section 1). Figure 1 provides a timeline highlighting major court cases and legislation that have influenced the education of language minority students since the ruling on Brown v. the Board of Education in 1954.
Figure 1. Legislation and Court Cases Affecting the Education of Language Minority Students
Theories of Second Language Acquisition

In 1982, Stephen Krashen proposed five hypotheses for second language acquisition: 1) the acquisition v. learning hypothesis, 2) the natural order hypothesis, 3) the monitor hypothesis, 4) the comprehensible input hypothesis, and 5) the affective filter hypothesis. For the past 30 years these hypotheses have formed the cornerstone for teacher training on language acquisition and program development for English language learners.

The first of Krashen’s (1982) hypotheses, the acquisition v. learning hypothesis, differentiates between language learning, a formal process of learning about the language, and language acquisition, an informal process of ‘picking up’ the language. According to the theory, for acquisition to occur, the student must be immersed in language within meaningful and comprehensible contexts. The emphasis here is on the “meaningful and comprehensible” component of the theory. In a more recent publication, Crawford and Krashen state, “When we can’t understand a language, we don’t acquire it. Incomprehensible input becomes undifferentiated noise, signifying nothing. It fails to register in the brain” (2007, p. 18-19).

Krashen’s comprehensible input hypothesis has been considered by many to be one of the most important elements of language acquisition. It builds upon the understanding that the language being learned must somehow be made comprehensible to the learner, which can be accomplished through the use of such things as visuals, realia, modeling, and gestures (Zainuddin, Yahya, Morales-Jones, & Ariza, 2002). A vital component of the input hypothesis is that the input needs to be slightly above the student’s current level of comprehension for learning to take place. This is commonly referred to as “i + 1”, comprehensible input plus one (Krashen, 1982). This concept proves to be true not only for language acquisition but for all other learning
as well. When input is at a level higher than “i +1”, the result is frustration and anxiety on the part of the student. When student anxiety is at such levels, learning is blocked.

**Comprehensible Input vs. Comprehensible Output**

Despite the wide acceptance of Krashen’s aforementioned five hypotheses, they have not gone without criticism. Krashen claimed that “we acquire language in only one way: when we understand messages in that language”, further stating that “comprehensible input is the only true cause of second language acquisition” (1984, p. 61). As it pertains to interaction and speaking, Krashen asserted, “Speaking is a result of acquisition and not its cause. Speech cannot be taught directly but ‘emerges’ on its own as a result of building competence via comprehensible input” (Krashen, 1985, in Gass, 1997, p. 82). These claims have come under particular scrutiny, with many scholars challenging Krashen’s claim that there is ‘only one way’. One such researcher has been Merrill Swain, introducing the “comprehensible output” hypothesis in 1985 to counter Krashen’s emphasis on “comprehensible input.”

Swain’s (1985) output hypothesis came out of her work with English speaking children in French immersion programs in Canada. What Swain found was that, while assessments of the listening and speaking skills of students in the immersion classrooms were consistent with their native French speaking peers, their speaking and writing skills lagged far behind. As Swain points out, “No one could possibly argue that immersion students did not receive an abundance of comprehensible input” (Swain, 2005, p. 472). Clearly, something more was needed to fully develop students’ language and increase their productive skills.

In the comprehensible output hypothesis, Swain (1985) contended that input alone is not sufficient in fully developing a second language. Directly contradicting Krashen’s assertion that
speaking is not a cause but a result of acquisition, Swain proposed that it is precisely through the production of either spoken or written output that language acquisition occurs.

According to Swain (1993), there are four ways in which output plays a role in the process of second language learning. First, language production provides the opportunity for meaningful practice and helps to develop automaticity or fluency. For many teachers, this is the obvious and primary reason for providing opportunities for interaction in the classroom. Second, output may force the learner to move from semantic processing, getting the meaning, to syntactic processing, adjustments to form, by helping learners to recognize gaps in their language and focus on reconciling those gaps. The third role of output is that of hypothesis testing, in which the learner tries out new expressions to see if they are using them correctly. The fourth and final way that Swain suggests output contributes to language learning is that it may prompt responses that can provide learners with information about the comprehensibility or accuracy of their utterances. Such responses may include confirmation checks, clarification requests, or corrections. With the information provided by the interlocutor, the learner can then modify or reprocess their output to make it more comprehensible. This fourth role for output is the focus on the notion of negotiation for meaning, essentially the back and forth process that interlocutors go through to arrive at a common understanding.

Negotiation for Meaning

Is miscommunication a bad thing? Should our goal, as teachers, be to eliminate miscommunication from our language learning classrooms? According to Gass and other interactionist theorists and researchers, the answer is no. Miscommunications can actually serve the purpose of developing students’ language and is, in fact, “beneficial from the standpoint of learning” (Gass, 1997, p. 104). Related to the previous discussion on comprehensible output,
*negotiation for meaning* is the reason why miscommunication should have a place in the language-learning classroom.

As defined by Gass (1997), the process of negotiation in language acquisition “refers to the communication in which participants attention is focused on resolving a communication problem as opposed to communication in which there is a free-flowing exchange of information” (p. 107). Pica (1994) tells us that the focus on interaction's role in the language learning process and negotiation for meaning, in particular, arose in the work of Hatch and Long during the late 1970s and early 1980s. In 1978, Hatch proposed that, rather than focusing on how the practice of previously taught language structures leads to more communicative use, research in the language learning process should consider the opposite relationship and investigate how communicative use could lead to the development of students’ syntactical understanding.

Long is one of many who took up Hatch’s challenge. In a study of negotiation, Long (1981) explored interactions between native speakers (NS) and non-native speakers (NNS) by first establishing a baseline with NS – NS interactions. Long then compared NS-NNS interactions, recording differences in the two types of pairings in their use of the interactional functions of expansion, repetition, and clarification in the negotiation process. From his findings, Long proposed the hypothesis that “participation in conversation with NS, made possible through the modification of interaction, is the necessary and sufficient condition for SLA” (p. 275). This has since been referred to and is often cited as Long’s “interaction hypothesis.”

One aspect of Long’s 1981 hypothesis required that interaction take place between a native speaker (NS) of the target language and the non-native speaker (NNS), language learner, with the assumption that it was necessary for the language learner to have the feedback of a
native speaker in order to benefit from the interaction. Fifteen years later, Long (1996) reiterated the need for interaction with a native speaker indicating, “Negotiation for meaning by definition involves … utterances by a competent speaker, such as repetitions, extensions, reformulations, rephrasings, expansions, and recasts” (p. 452). Swain’s early version of the output hypothesis, although not explicitly stated, seemed to support Long’s requirement that interactions be between native and non-native speakers. According to Swain (1985):

> Negotiating meaning needs to incorporate the notion of being pushed toward the delivery of a message that is not only conveyed, but that is conveyed precisely, coherently, and appropriately. Being ‘pushed’ in output… is a concept parallel to that of the i + 1 of comprehensible input. (in Swain, 2005, p. 472-473)

The use of the terms ‘precisely’ and ‘appropriately’ would imply the belief that a native speaker’s input in the interaction is a necessity of the negotiation process.

A study by Varonis and Gass (1985) challenged the claim that the interactions must be between native speakers (NS) and non-native speakers (NNS). The results of the Varonis and Gass study with college-aged students in an intensive English program indicated that “a greater amount of negotiation work takes place in NNS–NNS discourse than in either NS–NS or NNS–NS discourse” (p. 84). They suggest that interactions between native speakers and non-native speakers “actually discourages negotiation” because of the “inequality in the (language) status of the participants” (p. 86). Varonis and Gass attributed the higher negotiation rates between language learning peers to the fact that the individuals in the pairs were both non-native speakers and were more comfortable admitting and recognizing miscommunications because of their “shared incompetence” (p. 84). They argued that non-native speakers “do not lose face by
negotiating meaning [with other non-native speakers] in the same way they might with native speakers” (p. 85).

Mackey (2012) discusses similar findings from a 2003 study by Mackey, Oliver, and Leeman, which examined differences in adult and children’s interactional feedback in combinations of NS and NNS dyads. Mackey reported that, in the 24 child dyads, significantly more modified output occurred when the learners were interacting with other non-native speakers than when they were with native English speakers.

Negotiating the Negotiation Model

The Varonis and Gass 1985 study proposed a basic two-part model of negotiation called the “Trigger and Resolution model” (in Gass, 1997). In this model, the trigger is the stimulus that begins the negotiation process. The resolution that follows is the effort between the interlocutors to resolve the miscommunication. What occurs within the resolution phase is where the negotiation and the learning actually take place. This is certainly the most interesting part of the process and has subsequently been the focus of a great deal of research over the years. This subsequent research has served to expand and clarify the model, in essence “negotiating” the negotiation model.

Varonis and Gass (1985) articulated the back-and-forth nature of the model, using a very straightforward coding scheme in their study: T = Trigger, I = Indicator, R = Response, RR = Reaction to Response. Once a trigger (T) (misunderstanding) occurs, one of the conversants indicates a problem has occurred with some sort of indicator (I) or signal, marking that there has been a breakdown in the communication. After a problem has been indicated, there is a response (R), which is followed by a reaction to the response (RR). The following is a chant activity that I use with students that illustrates the negotiation model in a very simplified format:
(T) S1: This is a book.

(I) S2: A what?

(R) S1: A book.


S1: That’s right.

A study by Nakahama, Tyler, and Van Lier (2001) looked specifically at the trigger (T) component in the negotiation process as NS/NNS dyads participated in two different types of interaction activities: an open-conversation and an information gap. In analyzing the interactions, the researchers classified negotiation triggers into four types: 1) *Lexical* – verb/noun phrases and word choice; 2) *Morphosyntactic* – verb inflection, partitives, and plural morphemes; 3) *Pronunciation*; and 4) *Global*—discourse, content, or both (p. 384). The researchers found that the main trigger type used in the conversational activity was the global trigger. Whereas, in the information gap, lexical triggers were more prominent, with global triggers initiating only one-quarter of the negotiation cycles.

As we continue to “negotiate” the model, it is helpful to consider the role of communication strategies in the negotiation for meaning process. The use of communication strategies appears in the *Response* phase of the negotiation cycle. Tarone (2005) defines communication strategies as “mutual attempts of two interlocutors to agree on a meaning in situations where requisite meaning structures do not seem to be shared” (p. 488). In other words, communication strategies encompass the various ways that individuals might go about trying to convey meaning when a perceived breakdown has occurred. This is, of course, not limited to learners of a new language. All speakers experience loss of communication coherence even when conversing with speakers of the same language. When these breakdowns occur, or are
perceived to occur, we adjust. We naturally find ways to make our meaning known through a variety of communication strategies or redirect the exchange. The strategies chosen to deal with, or not deal with, the miscommunication can take on a variety of forms, influenced by interwoven factors that include culture, language fluency (or perceived fluency), and the context in which the miscommunication occurs.

As illustrated in Figure 2, Tarone (2005) classified communication strategies into two broad types: *Compensatory* and *Avoidance/Reduction*. With the Compensatory strategies, the individual tries to communicate their idea in one or more ways. Compensatory strategies can be placed into six categories. The first is through *Holistic Approximation*, which involves referring to a similar or related object (e.g. “It’s kinda like a…”). I find it helpful to consider this in terms of using an analogy to help fill in the gap. Two other strategies involve *Circumlocution* and are more analytical in nature. These involve describing the properties of the referent in terms of either its physical or functional properties. *Code-Related Strategies* can include: *Code-switching* (using the first language in place of the unknown word or phrase), *Foreignizing* (making a word from the first language sound like it fits the target language, usually done using accent or affixes), *Word Coinage* (inventing a new word either with compounding or affixes), *Non-Verbal Gesture*, *Sound Imitation* (onomatopoeia), and *Appeal for Assistance* (usually incorporates first language; e.g. “What’s the word for --- in English?”). An individual may use one or more of the strategies in a single negotiation cycle as the interlocutors work back-and-forth to achieve comprehension. Some strategies may even be used in combination. Non-verbal gestures, for instance, are often added to the other compensatory strategies to elaborate on the meaning.
Rather than try to repair a potential miscommunication, a speaker may choose to employ avoidance or reduction instead, which simply involves not referring to the object or idea at all (avoidance) or cutting short or redirecting the conversation (reduction). The reasons for avoidance or reduction can be varied, but may include lack of vocabulary, unfamiliarity or discomfort with the topic, or the speaker may perceive the effort needed to be fully understood not worth pursuing. Avoidance and reduction are more likely to occur in open-conversational activities than in task-based activities where a specific outcome is needed. Open conversation allows both speakers more latitude in that the individuals can more easily guide the conversation and control their language use (Long, 1996).

A similar avoidance or reduction strategy appears to occur in the writing assessments of English language learners (ELLs). Because receptive skills are generally believed to develop faster than productive skills, it’s an interesting anomaly when students’ writing scores exceed their reading scores on standardized tests. In Washington State, for example, 25.5% of fourth
grade ELLs met standard in writing on the 2010 statewide Measurement of Student Progress (MSP). In contrast, only 20.4% met standard in reading (Washington State Office of the Superintendent of Public Instruction, 2010). ELL teachers and the students themselves have noticed a similar trend in individual performance on the Washington Language Proficiency Test (WLPT) for some time and attribute it to students having more control over their writing, because they can “choose the words”, as one student told me. A future study along these lines might prove enlightening.

Returning to the final phase of the negotiation sequence, the Reaction serves either to indicate that a repair is complete and the interaction can move on, or it can serve as a re-signal, indicating that additional strategies should be employed to repair the miscommunication. Reactions may resemble the indicator or signal in many respects and can take the form of non-verbal gestures, pragmatic markers, or a variety of transitions signaling the conclusion of the particular negotiation cycle. A single negotiation sequence may only take a brief moment, involving merely a look of confusion, replacement of a word, and acknowledgment. Alternatively, the sequence may go on for several seconds as the conversants move through the Response – Reaction cycle for several rounds before the meaning becomes clear and the participants have successfully negotiated for meaning and arrived at resolution.

Figure 3 brings together the various components that have contributed to negotiating the negotiation model thus far and emphasizes the back-and-forth nature within the resolution phase of the cycle.
It is, of course, possible that the meaning may not be fully negotiated, in which case the participants may choose to abandon the negotiation and move past the difficulty. This ‘giving up’ on the negotiation is more likely to occur in open-ended tasks than in closed tasks. As Long (1996) observed, “free conversation is notoriously poor as a context for driving IL [Inter-Language] development.” He concluded that, “the lack of any fixed topics or outcomes permits rapid, superficial treatment of topics and the dropping of any that cause linguistic trouble” (p. 448).

The type of interaction in which the language learners engage can certainly alter the amount and type of negotiation that takes place. Again referring to Long (1996):

Tasks that orient participants to shared goals and involve them in some work or activity produce more negotiation work. The nature of the [two-way closed] task causes topics...
and subtopics to be recycled until solutions are reached, producing more negotiation work. (p. 448)

The previously referenced study of Nakahama, Tyler, & Van Lie (2001) also examined frequency of negotiation repair and complexity of interaction as they compared the task-based and open-conversational activities. Their findings supported Long’s conclusion. In the NS/NNS dyads, the task-based activity, in the form of an information gap, prompted significantly more instances for negotiation repair than the less focused open-conversational activity. However, upon investigating interactions more qualitatively, the researchers found that, although negotiation was less frequent in the open conversation activities, the utterances of the non-native speakers (NNS) were much longer and more complex than during the information gap activity. However, there was not a comparable difference for the native speakers (NS) in the dyads.

It has been my experience, as a teacher practitioner, that there is a place for both types of interaction in the language-learning classroom. Whether students are working together to solve a particular problem, accomplish a specific task, or are engaging in open conversation with native or non-native speakers, their language is being “pushed” in different ways, providing different benefits in the language learning process.

Building on Krashen’s (1982) “monitor hypothesis” and nested in the interactionist paradigm of Long (1996), Gass (1997), and Swain & Lapkin (1995), Chapelle (1998) provided a summary of the language acquisition process as described by Gass (1997). The process begins with the learner being exposed to some sort of input—either visual or auditory. As the learner is able to apperceive or notice aspects of the input that are presented in the target language, only those elements of the input that are noticed can pass to the comprehension level where it may be able to be processed. It is important to note that, according to this model, only elements that
have been made salient to the learner in the *apperception* phase will be processed in the *comprehension* phase. Within the comprehension phase, when semantic and syntactic processing are combined, input becomes comprehended *intake* that can then become *integrated* into the learner’s short term memory to become part of their developing linguistic system. Figure 4 provides a summary of the model.

![Diagram](image)

*Figure 4. Input, comprehension, output process. (adapted from Chapelle, 1998)*

According to Chapelle, the *output* phase develops as a result of the integration process and provides observable evidence of the process as well as a means to linguistic development. Although this would appear to be the endpoint of the discussion, when a learner is able to produce output and interact in the target language, interactionists contend and Chapelle (1998) reminds us, the real process of acquisition is only beginning to take place.

**A Shifting Paradigm?**

Work done over the past 25 years in the area of interaction, specifically in negotiation for meaning, has sought to move away from Krashen’s emphasis on comprehensible input as “the only true cause of second language acquisition” (1984, p. 61). The literature is, however, rife with statements on how interaction and negotiation for meaning serve to increase the comprehensibility of input (e.g. Chapelle, 1998; Long, 1981, 1996; Nakahama, Tyler, & Van Lier, 2001; Pica, 1994). Even Swain’s own work, which clearly set out to counter Krashen’s emphasis on comprehensible input, is still filled with evidence and arguments for making input comprehensible (Swain, 1985, 1993; Swain & Lapkin, 1995). Instructional and teacher training
models developed in the 1980s and 1990s, such as CALLA, SIOP, and GLAD, have emphasized the importance of making content comprehensible. Thus, even though comprehensible input may not be ‘sufficient’, nor the ‘one and only way’ that students learn in their new language, it certainly is a major player. As Pica (1994) articulated, “comprehensible input, however modified, might not be efficient, or even sufficient, for SLA” (p. 274). For students to successfully acquire a new language, they need not only to understand or comprehend the input, but additionally they need to be engaged with sufficient and appropriate opportunities to interact in the target language.

Whether a result of a shifting paradigm, change in communicative academic style, or well-garnered wisdom, seminal authors have reframed their original theories and hypotheses from something of an absolute “all-or-nothing” stance to a more interpretivist or constructivist view with multiple possibilities and an emphasis on process. As previously mentioned, Long (1981,1996), adjusted his interaction hypothesis over time. As he puts it 15 years after its introduction:

The updated version of the Interaction Hypothesis involves a mix of well and less well-established L1 and L2 acquisition research findings, some rather high inference interpretation, and some speculation. It is certainly not intended, of course, as anything like a complete theory of language learning. (1996, p. 453)

Gone, it would seem, are the absolute statements of ‘necessary’ and ‘sufficient’, using instead phrases like ‘contribute to’ and modals such as ‘may’ and ‘should’.

Swain (2005) has also adjusted her hypothesis, re-articulating her “output hypothesis” to emphasize process over product (Notice the intentional dropping of the word ‘comprehensible’ in her title.). Swain explicitly places a greater emphasis on cognitive processing and
metalinguistic function in the updated hypothesis. References to Vygotsky’s sociocultural theory newly appear in the updated hypothesis, as do quotes by Smagorinsky (1998), the following of which, I believe captures well this re-framing: “The process of rendering thinking into speech is not simply a matter of memory retrieval, but a process through which thinking reaches a new level of articulation” (in Swain, 2005, p. 479).

As interaction and negotiation continued to be negotiated in the literature, Swain (2005) reflected on the last-quarter century and her new thinking:

Looking back over the research of the late 1980s and 1990s, I think the label “comprehensible output” tended to get in the way of the idea of output as process even though the 1985 formulation was clear that the output hypothesis is about what learners did when pushed, what processes they engage in…The label (noun) put the focus on product, rather than on process. Much of the research conducted in the late 1980s and 1990s about the output hypothesis was descriptive in nature, and tended to focus more on occurrence and acquisition. (p. 473)

The noted absence of research on the process and situationality of interaction, in-line with the theories of Vygotsky and the work of Smagorinsky, are indeed notably absent in this period. Tarone (2005) noted similarly that research on communication strategies “has not maintained the original emphasis on the essential contextuality of communication strategy use” (p. 489). Tarone and others call for research using various methods of data collection, with “less emphasis on taxonomic description and classification, and more emphasis on looking at the learner’s communication strategy choices in the context of situation” (ibid).
The Role of Anxiety and The Affective Brick Wall

The affective domain has long been considered of importance in learning. The social and emotional aspects may be of even greater importance in language learning and are therefore addressed specifically in the “affective filter” hypothesis as introduced by Dulay and Burt in 1977 (Krashen, 1981). Following the work of Gardner and Lambert in 1972, Stevick in 1976, and Dulay and Burt in 1977, Krashen (1982) identified three affective variables that influence second language acquisition: 1) anxiety, 2) motivation, and 3) self-confidence (p. 56). The affective filter hypothesis contends that when students are in a highly stressful environment their anxiety levels increase, thus raising the affective filter, and blocking input. (My own experience of teaching a driver’s education course for English language learners would support this. However, I conceive of it as an “affective brick wall” rather than an “affective filter”.) Krashen stresses that students acquire language best in low anxiety environments, where they are free to take risks with the language without being ridiculed. As Crawford and Krashen (2007) so aptly state:

The sink-or-swim treatment actually retards the process of English acquisition, as well as academic achievement. Failure rates for children who were subjected to this approach (before it was outlawed) were extremely high. For those who survived, a common recollection is…no recollection of anything that happened in the first two to three years of school – that is, until they began to acquire some English. Nothing comprehended, nothing learned. (p. 19)

Sheltering and Instructional Strategies

Throughout the United States, more and more school districts are relying on classroom teachers to serve the needs of a rapidly growing population of students designated as Limited
English Proficient (LEP) while providing them with little training to do so. According to Crawford and Krashen (2007), the average classroom teacher working with ELLs has had only 4 hours of ELL training over the past 5 years. This amounts to only 48 minutes per year. Given the rapid increase in the number of English language learners in our schools, an increase of 241% in Washington State since 1990, this is clearly not enough. In the past few years, partially in response to mandates from No Child Left Behind in 2001, districts and individual schools have been providing increased training for their instructional staff working with English language learners.

According to Washington Administrative Code, WAC 181-78A-270 (a) (i), candidates in teacher preparation programs are required to demonstrate competence in “using multiple instructional strategies…to address student academic language ability levels and cultural and linguistic backgrounds.” Additionally, WAC 392-160-010 (3) requires that districts provide professional development training on the “appropriate use of instructional strategies” and on “instructional materials for use with culturally and linguistically diverse students.” Although Washington State does not specify any particular instructional programs to be used with English language learners, the Washington State Transitional Bilingual Instruction Program (TBIP) (2009) guidelines indicate that such training should prepare teachers and other instructional staff to be able to implement:

- Strategies that integrate language acquisition and academic achievement at the same time
- Strategies that promote proficiency in English (and the primary language, where applicable) for academic purposes, including literacy
- Strategies such as SI (Sheltered Instruction) that ensure that academic instruction
through the second language is meaningful and comprehensible to second language learners. (p. 25)

So, what types of training programs are currently being implemented to meet these goals? In the 2010-2011 annual report to the legislature, the Washington TBIP (Malagon, McCold, & Hernandez, 2011b) specifically identified three programs used in Washington for implementing strategies into the classroom: *Guided Language Acquisition Design (GLAD)*, *Sheltered Instruction Observation Protocol (SIOP)*, and *Cognitive Academic Language Learning Approach (CALLA)*. Before moving forward with a discussion of what these particular training models involve and the underlying strategies that they employ, I would like to first elucidate what is meant by the term *sheltered instruction*, a term which has undergone considerable change over the years.

**What Was and Is “Sheltered Instruction”?**

As it was first introduced during the 1980s, *sheltered instruction* or *sheltered English*, also known as *SDAIE* (Specially Designed Academic Instruction in English), was an “instructional approach used to make academic instruction in English understandable to LEP students” (Freeman & Freeman, 1988, p. 2). The term referred to classes in which Limited English Proficient (LEP) students *who already possessed some English proficiency* received specially designed content area instruction outside of the mainstream classroom. According to Freeman and Freeman, “Students in these classes are ‘sheltered’ in that they do not compete academically with native English speakers since the class includes only LEP students” (p. 2). The goal was to “mainstream the student gradually” (p. 3).
As is seen in a side-by-side comparison of Washington State 2007 and 2009 program descriptions as shown in Table 1, although the “sheltered” terminology has become more dominant, the concept of ‘sheltering’ students in the sense of Freeman and Freeman is gradually falling by the wayside. Also notice that “content-based ESL” is used interchangeably with “sheltered instruction,” a point that will be addressed later in this paper.

Table 1

*Comparison of Previous to Current TBIP Definition*

<table>
<thead>
<tr>
<th>2007</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Content-Based ESL/Sheltered Instruction:</strong> Students are taught entirely in English through ESL techniques. Teachers must develop the students’ academic language proficiency consistently and regularly as part of the lessons and units that they plan and deliver. Teachers generally present the regular, grade-level subject curriculum to the students through modified instruction in English, although some special curricula may be designed for students with significant gaps in their educational backgrounds or very low literacy skills. This model requires the teacher to have significant training in second language acquisition strategies as well as support for effective implementation. (TBIP, 2007, p. 12)</td>
<td><strong>Sheltered Instruction (SI)</strong> is an approach used widely for teaching language and content to ELLs, particularly as schools prepare students to achieve high academic standards. In SI, academic subjects (e.g., science, social studies) are taught using English as the medium of instruction. SI is most often used in classes comprised solely of ELLs, although it may be used in classes with both native English speakers and ELLs when necessitated by scheduling considerations or by small numbers of ELLs. (TBIP, 2009, p. 24)</td>
</tr>
</tbody>
</table>

The Washington State TBIP guidelines indicate that Sheltered Instruction (SI) is for “students with *limited or no English*” (2009, p. 26) [emphasis added]. This is a marked shift
from Freeman and Freeman’s (1988) description of SI, which indicated that sheltered instruction is intended for “ESL students who already have some English proficiency” (p.4) [emphasis added]. It is also noteworthy that a definition for English as a Second Language (ESL), while included in the 2007 program guidelines, does not appear in any form in the newer 2009 version. The 2010-2011 report to the legislature indicated that “ELLs at the elementary level, who receive push-in support from TBIP staff or limited pull-out in the context of a basic education grade-level classroom, would also fall under this (sheltered instruction) model” (Malagon, McCold, & Hernandez, 2011b, p. 19). The loss of the ESL Pullout designation as a separate category in Washington State impacts how one should interpret state and federal reports. One will no doubt see an increase in numbers reported for “Sheltered Instruction” models when no such real increase exists; student numbers have merely been shifted from one category to another. It is important to bear this shift in mind as one looks at reports of program effectiveness utilizing this new designation that pools ESL pull-out and push-in with SI.

**Overview of Three Sheltered Models**

Based in a constructivist paradigm, which posits that learning is an actively constructive process and that new information is linked to prior knowledge (Learning Theories Knowledgebase, 2011), all three of the sheltered instruction/content training models identified in Washington State have strong roots in Krashen’s comprehensible input hypothesis, cognitive learning theory, and Vygotsky’s social development theory, which argues that social interaction precedes cognitive development. According to Vygotsky, the internalization of learning comes as a result of socialization and social behavior. The work of Krashen, Cummins, Piaget, Vygotsky, and Bruner figure prominently in the theoretical framework of all three models. As we move forward, I will first provide an overview of each model, taking them in chronological
order, and finish by highlighting the common underlying strategies upon which they are based.

**Cognitive Academic Language Learning Approach (CALLA)**

*CALLA is a research-based instructional program that fosters the school achievement of students who are learning through the medium of a second language. This professional development program focuses on science and math.* (TBIP, 2011, p. 13)

The earliest of the three models, introduced by Chamot and O’Malley in the mid 1980s, the Cognitive Academic Language Learning Approach (CALLA) is a “content-based curriculum to prepare upper elementary and secondary LEP students for a transition to the mainstream subject areas” (1986, p. 6). The point is stressed throughout the literature that CALLA is meant to be a “transition” or “bridge” to the mainstream or grade level classroom (Chamot & O’Malley, 1986; Chamot & O'Malley, 1994; Orem & Wang, 1997). Chamot and O’Malley (1986) explicitly state, “it is not an immersion program, and it does not substitute for either the ESL or the mainstream program” (p. 6).

As a content-based program, as opposed to a program of *sheltered instruction*, the primary emphasis for CALLA is on developing the students’ language, specifically their *cognitive academic language proficiency*, as coined by Cummins. While often used interchangeably, as seen in the previous program descriptions, there is, in fact, a distinction between the terms. Undoubtedly a result of the mandates of NCLB in 2001, *Sheltered Instruction*, by contrast, places content learning first with a secondary focus on language learning (Chamot & O'Malley, 1987; Zainuddin, Yahya, Morales-Jones, & Ariza, 2002). The underlying theories and basic strategies remain very similar regardless of the changing label.
According to the program’s authors, CALLA is rooted, as are many of the programs coming out of the educational reform movement of the 1980s and 90s, in the theories of educational psychologists Anderson, Ausubel, Bruner, Piaget, and Vygotsky (Chamot & O'Malley, 1994). The fundamental principle is that language is not only used to illustrate or explain what is learned, but actually facilitates the thinking process and contributes to the construction of knowledge within the learner. Chamot and O’Malley included an illustration of this relationship in the form of a “Cognitive Model of Learning” (p. 20) tree with the previously mentioned “roots”. CALLA is shown as one branch of the tree, alongside similar programs: Language Across the Curriculum, Language Experience Approach, Whole Language, Process Writing, Cooperative Learning, and Cognitive Instruction. Although these were seen as separate programs when the CALLA handbook was published in 1994, several have since been combined and have become part of several newer programs developed to meet the needs of English language learners.

CALLA consists of three components: 1) a content-based curriculum correlated to the mainstream program; 2) English language development specific to the academic areas of science, mathematics, and social studies; and 3) instruction in the use of learning strategies (Chamot & O'Malley, 1986).

The first component indicates that the content taught should include the same content topics that students will encounter when they move into the grade-level classroom (Chamot & O'Malley, 1994). By aligning the content with what is also being taught in the mainstream classroom, students gain confidence and increased motivation because they are learning ‘real stuff’. The recommendation of the CALLA approach is to introduce new subjects gradually rather than all at once, so as not to overwhelm learners with new content and new language
simultaneously. What many teachers find surprising is that CALLA places science as one of the first content areas to be introduced despite the high level vocabulary of the subject area. The rationale is that the context-rich learning environment of an investigative science curriculum along with hands-on experiences in a science classroom help foster the academic language learning that is the goal of the CALLA model.

In the second component, as teachers begin planning their courses, it is necessary for them to analyze the language demands of the different content subjects being used. This should include both the language of the curriculum materials as well as the language needed for classroom participation in that content area. The students would then need explicit instruction in the actual language functions, structures, and subject specific vocabulary that they will need as they transition into the mainstream classrooms.

In chapter 2 of her book, *Learning to Learn in a Second Language*, Pauline Gibbons (Gibbons, 1991/1993) has taken the planning for language learning concept to the lesson/activity level and provides an excellent list of language functions with descriptions along with a language planning framework to help teachers plan for the language functions, language structures, and specific vocabulary that students need to fully engage in the given activities of the lesson. Herrell and Jordan (Herrell & Jordan, 2008) have adapted Gibbons’ framework and included it in their text, *50 Strategies for Teaching English Language Learners*. I have found it useful to modify the planning sheet slightly by moving the vocabulary needs to the beginning of the framework. Most pre-service and in-service teachers with whom I’ve worked have found it easier to think about the language functions and structures by beginning with the vocabulary needs first rather than placing them last. I have also found that most of the pre-service teachers, as well as some in-service teachers, have a tendency to focus only on the receptive language needs of the lesson,
considering only the input and what the students will need to read or hear to comprehend the content. Even with prompting, they seem to find difficulty in crossing over into considering the output needs of the lesson as well.

The third component, the explicit teaching of learning strategies, is based heavily on cognitive learning theory and supported by previous research (O'Malley, Chamot, Stewner-Manzanares, Russo, & Kupper, 1985; Chamot & O'Malley, 1986). The three types of learning strategies identified in the CALLA model are Metacognitive, Cognitive, and Social-affective.

With the metacognitive strategies, students actively plan and think about their own learning by self-monitoring both comprehension and production and by self-evaluating how well they have achieved the learning objective. One example might be selective-attention strategies, which include paying particular attention to various linguistic markers, such as signal words and phrases, which facilitate comprehension by indicating the type of information that is being expressed.

In cognitive strategies, the student transforms the material to be learned by manipulating it either mentally or physically. Two types of cognitive learning strategies identified by Chamot and O’Malley (1986) are transfer and elaboration. According to Chamot and O’Malley:

Transfer can be used to increase comprehension by deliberately remembering what is already known about the topic, and by capitalizing upon any linguistic similarities that may exist between the first and second languages to assist comprehension. In elaboration students consciously interrelate concepts into new information or integrate new concepts into their existing knowledge structure. (p.18)

The social–affective strategies require student interaction either cooperatively or by asking questions for clarification in order to facilitate the learning. Cooperative learning in pairs
or small groups is identified as an example of a social-affective strategy employed to “allay anxiety” and to “practice using language skill” (Chamot & O'Malley, 1986, p. 18). Even though both of these purposes for interaction in the language learning classroom are completely appropriate and necessary, it is surprising that the important role that cooperation and interaction play in constructing knowledge (Vygotsky & Luria, 1934/1994) is not identified in relation to social-affective strategies in either the 1986 report or article, nor in the 1994 handbook. Nevertheless, it is clearly a part of the theoretical framework (Chamot & O'Malley, 1986; 1994).

As we move on to the next training model, it is indeed interesting to note that, despite the emphasis Chamot and O’Malley place on the explicit teaching of learning strategies with English language learners in the CALLA model, they “have come to the conclusion that there are probably no unique strategies for learning a second language, although a subset of general learning strategies may be of particular use in language tasks” (1986, p. 17).

**Sheltered Instruction Observation Protocol (SIOP)**

*The Sheltered Instruction Observation Protocol is both a valid and reliable, research-based observation instrument as well as a guide for planning instruction that focuses on both the academic and linguistic needs of ELLs. The model is built on the premise that teachers possessing these skills will be prepared to provide English language learners with a better learning environment. Critical features of high quality instruction for English language learners are embedded within the SIOP model.* (TBIP, 2011, p. 12)

Originally developed in the early 1990s as an instrument to evaluate how effectively teachers implemented sheltered instruction (SI) in their classrooms, by 2002, SIOP had been developed into the SIOP Model for professional development (Echevarria, Vogt, & Short, 2008).
The program’s developers emphasize that the SIOP model is a sheltered instruction (SI) approach rather than a content-based ESL approach, indicating that SI emphasizes “grade level objectives delivered through modified instruction that makes the information comprehensible to the students” (p. 13). SIOP developers further claim that content-based ESL instruction (e.g. the CALLA model previously described) “has not been sufficient to help all ELs succeed academically” (ibid).

The SIOP evaluative tool originally consisted of three sections: Preparation, Instruction, and Review/Evaluation (Short & Echevarria, 1999, p. 13). The observation protocol is currently divided into eight components: Lesson Preparation, Building Background, Comprehensible Input, Strategies, Interaction, Practice/Application, Lesson Delivery, and Review/Assessment. The 30 items of the evaluative tool are divided among these eight components. Teachers who are observed with this protocol are evaluated on each item by way of a four-point Likert style scoring system. The SIOP Model for teacher training came about as teachers commented on the need to understand the criteria on which they were to be evaluated. By understanding the components of quality SI instruction, they could then incorporate these practices into their teaching.

The strategies utilized in the SIOP Model are based strongly in the theories of Vygotsky, Krashen, Bloom, and Chamot. The strategies can be broken down into three broad categories: Scaffolding Techniques, Higher-Order Questioning, and Learning Strategies.

Vygotsky’s Zone of Proximal Development (ZPD), essentially the difference between what a learner is capable of on his/her own and what he/she can do with assistance, is the basis for the concept of scaffolding described in the SIOP Model. The model can be separated into three types: verbal scaffolding, procedural scaffolding, and instructional scaffolding. Verbal
scaffolding uses prompting, questioning, and elaboration to facilitate comprehension and learning. Echevarria, Vogt, and Short (2008) identify the following forms of verbal scaffolding employed in the SIOP Model: paraphrasing, ‘think-alouds’, reinforcing contextual definitions, slowing speech, increasing pauses, and speaking in phrases. Examples of the second type, procedural scaffolding, includes explicit teaching, modeling, practice opportunities with others, independent application, one-on-one teaching, small group instruction, practicing with a more experienced student, and partnering/grouping for reading with more experienced readers. The third form, instructional scaffolding, could include the use of graphic organizers by the teacher as a pre-reading or organizational tool to clarify key ideas from the text. SIOP encourages the Teach—Model—Practice—Apply format for instructional scaffolding. This model for scaffolding instruction is reminiscent of Madeline Hunter’s popular Instructional Theory Into Practice (ITIP) model for lesson design from the 1980s upon which my first lessons were based, emphasizing objectives, anticipatory set, modeling, guided practice, and independent practice. The bottom line in regard to scaffolding is that lessons should be structured with a gradual release of responsibility, moving students increasingly toward independence.

Krashen’s comprehensible input hypothesis is as central to SIOP as it is to SI and Content-Based Instruction (CBI) models generally. One might in fact surmise that it is essential to instruction with native speakers in the mainstream classroom as well. In general terms, to make input comprehensible, effective teachers should:

- Use speech appropriate for students’ proficiency level (e.g., slower rate, enunciation, and simple sentence structure for beginners).
• Explain academic tasks clearly. Use a variety of techniques to make content concepts clear (e.g., modeling, visuals, hands-on activities, demonstrations, gestures, body language). (Short & Echevarria, 1999, p. 20)

One strategy recommended for making input comprehensible for English language learners is adapting text and the use of Hi-lo readers. In both cases, the language structures and/or vocabulary are adjusted to the student’s current level of comprehension in order to clarify the content and make the input comprehensible for the student. This strategy is identified throughout the SIOP Model (Echevarria, Vogt, & Short, 2008). At the same time, the developers of SIOP criticize the use of adapted text in the content-based programs, stating:

We have learned that if students’ exposure to content concepts is limited by vocabulary-controlled materials, the amount of information they learn over time is considerably less than that of their peers who use grade-level texts…instead of closing the gap between native English speakers and ELs, the learning gap is increased and eventually it becomes nearly impossible to close. (p.24)

The explicit teaching of learning strategies, as introduced in the discussion on CALLA, is considered a central component for the effective teaching of English language learners with the SIOP model as well. Metacognitive, Cognitive, and Social/Affective strategies are once again highlighted. Metacognitive strategies involve, essentially, thinking about your own thinking through awareness, reflection, and interaction. In practice, teachers are expected to have content and language objectives “clearly defined, displayed, and reviewed with students” (Echevarria, Vogt, & Short, 2008, p. 228). Learning logs, whereby students reflect on their learning and self-
assess how they are meeting the stated lesson objects, are examples of what is being called “student voice”.

Cognitive strategies, as described earlier with CALLA, actively involve students in organizing input from the curriculum, either mentally or physically, to enhance their understanding of the material. Some of the techniques associated with the cognitive strategies in the SIOP Model include: clustering, graphic organizers, semantic maps, outlines, and note-taking. The cognitive strategies are probably the most salient of the strategies in SIOP or any of the models. Although the organization and transformation of material can occur mentally, we often see evidence in the classroom in the form of physical products. In fact, the majority of cognitive strategies are described in physical terms, e.g. underlining, graphic organizers, two-column notes, outlining, clustering, etc. While developing understanding through mental processes and reasoning occurs inside the head, as the term implies, it is and will likely continue to be, primarily determined and evidenced through physical means.

The third group of learning strategies, the Social/Affective strategies, mirror what was indicated in the CALLA model in that they are primarily to clarify information, provide practice, and serve as a problem solving strategy. Although group work and partnering is encouraged throughout the SIOP model, it is unclear how a teacher goes about ‘explicitly teaching’, as the model requires, these particular strategies. Interaction and groups are clearly identified as important to promote student learning. As Echevarria, Vogt, and Short (2008) state, “teachers must create ample opportunities to practice using academic language, not simply social uses of language” (p. 115). Benefits of student interaction include: brain stimulation in the form of “pleasure structures” (p. 116), increased motivation, reduced risk, more processing time, and increased attention. SIOP developers point out that opportunities for interaction that support the
social/affective strategies are rare in the mainstream classroom and therefore must be promoted in SI models.

There seems to be considerable overlap between what constitutes scaffolding, comprehensible input, and interaction. As I hope I have made clear, much of what is defined in SIOP as effective teaching for sheltered instruction has strong parallels to the CALLA model: language and content objectives interrelated, an emphasis on developing academic language, opportunities for interaction, and the explicit teaching of learning strategies. Where SIOP differs is in the evaluative tool and lesson planning frameworks that help teachers determine 1) how their teaching compares to identified effective SI criteria; and 2) how to plan and implement lessons that include these effective strategies, thus promoting student learning in the content areas.

**Guided Language Acquisition Design (GLAD)**

*Project GLAD training provides research-based theory and practical, effective strategies for the development of academic language, literacy, academic achievement and cross-cultural skills of ELLs. Districts throughout Washington State have supported cohorts of teachers to complete GLAD training as well as investing in key trainers at the district level to offer ongoing GLAD training and support.* (TBIP, 2011, p. 12)

Originating with two ESL teachers in the Fountain Valley School District in Orange County, California during the mid 1980s, GLAD was first known as GLEP (Guided Language Experience Process). It was changed to GLAD (Guided Language Acquisition Design) in 1992 (Brechtel, 2001). GLAD is described as “a model of professional development in the area of language acquisition and literacy” (Project GLAD, 2010, n.p.).
Beginning with the works of Krashen and Cummins, Brechtel (2001) identifies the following influences on the development of GLAD:

- Language is acquired by meaningful immersion, demonstration, and opportunities to practice (Krashen, Cambourne).
- Comprehensible input (Krashen, Cummins)
- Low affective filter—high self-esteem, low anxiety, and inclusion of all students (Krashen, Cummins, Baron, Sagor, and Wink)
- Negotiation for meaning, comprehensible output, guided oral practice, zone of proximal development and scaffolding with emphasis on peer interaction (Long, Swain, Cummins, Vygotsky)
- Cognitively demanding, complex concepts and language (Shefelbine, Collier/Thomas)
- Teacher- and student-generated text (Brechtel/Haley, Van Allen)
- Brain research—patterning and the importance of the growth of dendrites and axioms connecting brain cells through interactive, kinesthetic learning and conversation (Kovalik and others)
- Multiple intelligences (Gardner)

The presentation of GLAD in this dissertation will be significantly briefer than the previous discussion of CALLA and SIOP because many of the underlying theories and strategies upon which GLAD is based have already been discussed in the other two models. There are, however, a few big differences. The first of which I will illustrate with the following reflection:

Upon attending my first day of GLAD training in January 2006, my first thoughts were that the strategies that they were introducing were really nothing new. I had been
teaching English language learners for almost 20 years at that time and saw many of the strategies as being the same old stuff in a new package. Many of my fellow ELD specialists tended to agree with me. The use of graphic organizers, charts, music, total physical response (TPR), and student interaction were all things which we already knew to be highly effective and which came quite naturally to the experienced ESL teacher. It would be interesting to see how this training, targeted primarily to classroom teachers without ELL background, would be different from the other trainings we had attended.

During the second day of the two-day background training, I began to see some of the theoretical differences, primarily in GLAD’s emphasis on the closely related areas of brain research, negotiation for meaning, and comprehensible output. And by the time I was into the weeklong observation of a fifth-grade class a month later, the application differences had become clear. The differences were not so much in the strategies themselves, but in the management of the classroom.

Almost without fail, when I have introduced these same strategies to classroom teachers, the question has come up, “That sounds great. But what do I do with the rest of the class?” What the classroom teachers saw in the five-day observation component of the training was how to incorporate the strategies, beneficial to language learners, into instruction for the entire class. The training also emphasized how to utilize small group instruction while maintaining engagement for the remainder of students not directly involved in the small group lesson through table tasks. The table tasks, worked on by each table group when the teacher is working directly with small groups, emphasize collaborative learning and individual responsibility to the table group.
GLAD makes explicit the connections between brain-compatible teaching and strategies. For example, in the section of the training on brain imprinting, Brechtel (2001) points out that teachers should “use graphic organizers, such as the pictorial input chart, that imprint information in a pattern that makes learning retrievable over a long period of time” (p. 11). Rather than merely telling teachers that they ‘should use graphic organizers’ or making general statements about ‘helping students to organize their thinking,’ GLAD helps teachers understand why the strategies work, making it more likely that they will buy-in to the strategies and actually use them when they are back in the classroom.

The same sorts of explanations are applied to Long’s negotiation for meaning and Swain’s comprehensible output theories. In GLAD, the purposes for interaction are identified beyond reducing anxiety and providing practice, as they are in CALLA and SIOP. Group work and interaction are identified as the mechanisms that actually cause the learning to take place. It is emphasized that through the social interaction of processing the language, that acquisition takes place, a point made clear in the work of Swain (1985, 1993, 2005), Long (1996), and Vygotsky & Luria (1934/1994).

In the areas of reading and writing, GLAD emphasizes the need for scaffolding that moves from 1) teacher modeling for the whole class to 2) small group (cooperative) practice of what has been modeled to 3) individual use of what has been previously model by the teacher and then practiced in groups (Brechtel, 2001). Some strategies for reading and writing include DRTA (Directed Reading Thinking Activities) and SQ3R (Survey, Question, Read, Recite, and Review), cooperative strip paragraph, poetry frames, team poetry, chants, interactive journals, etc. Many of these strategies are already well know by classroom teachers, appearing in other sheltered, content, or literacy programs in the mainstream. What is different in the GLAD model
is the emphasis on scaffolding, from whole group → small group → individual and the strong focus on interaction.

Student-created texts are a vital part of the GLAD model. As Brechtel points out, “The walls should be dripping with the language of the students and the content they are learning” (2001, p. 15). In this same vein, it is recommended that:

- Classroom walls are no longer filled with merely cute little art. Rather, the art is combined with the print process, inquiry charts, input charts, and others on a daily basis.
- Read and write on the walls often—we call this focused reading. (p. 17)

For my own classroom, I like to start with almost empty walls and have the students create what goes up. For the ELL students, when they help create it, they know what it means and are reminded of the moment and process each time they see what is on the walls. In this way, neurologic pathways are strengthened, language is mentally rehearsed, and learning takes place. For my middle school, high school, and even adult students this also helps to create classroom community. There’s a sense that we are all in this together.

There are literally hundreds of strategies or techniques that can be employed for effective teaching in today’s linguistically and culturally diverse classrooms. A few of the basics, common to the three models, include: integrating language skills with content, use of visuals, advance organizers, hands-on experiences, graphic organizers, scaffolding, think-alouds, modeling, balance of direct instruction with flexible grouping, and the many forms of interaction. (See the Appendix A for a Wordle overview of the strategies related to each model.)

Specific strategies for effective teaching are in fact limitless as the teacher takes a known strategy and adapts it to the needs of his/her students. In the end, I think Chamot and O’Malley
may be right in that “there are probably no unique strategies for learning a second language” (1986, p. 17). As my former colleague, Susan Stannard, often stated “It’s just good teaching!” But, are good teaching and strategies enough? Or is the environment in which these strategies are employed and the peers with whom the students interact of equal if not more importance?

Into The K-12 Classroom: What the Research Says

A 2008 report by the U.S. Department of Education, A Nation Accountable, indicated, “only about 6 percent of English language learners are proficient in reading” (p. 13). Although this is an often-quoted statistic, it can be misleading. An English language learner is defined by the U.S. Department of Education as any student who has “sufficient difficulty speaking, reading, writing, or understanding the English language to deny such individual the opportunity to learn successfully in classrooms where the language of instruction is English” (United States, 2010). Therefore students are identified as ‘ELL’ only if they are not yet proficient in reading and writing in English. An ELL is by definition not yet successful. Once successful students achieve proficiency, they are no longer designated as ‘ELL’ on high stakes assessments in subsequent years. Therefore, scores for ELLs are only counted as proficient the first year that they achieve benchmark, causing the reported achievement for students in this subgroup to be forever low. In most cases, the scores of successful ELLs are no longer included in the reporting on the ELL subgroup; the successful scores are moved into the comparison group and actually weigh against the scores of the remaining ELL group. Because students move out of the ELL designation once they attain standard, they are not a true demographic. As Thomas and Collier point out, “ELLs do not remain ELLs forever” (2002, p. 24).
Program Models and Reading Achievement

The learning environment, which includes the type of program model, is, not surprisingly, an important consideration in language learners’ achievement. An extensive five-year study by Thomas and Collier (2002) followed the achievement of language minority students in sixteen school districts across the United States. This longitudinal study showed a wide variety of ELL achievement, indicating a primary factor for the differences to be the type of instructional program employed by the participating districts. Students who were mainstreamed immediately and received no additional language support (parents waived ELL services) showed the lowest reading achievement, with first grade students at the 48th NCE (Normal Curve Equivalent) during the first year of the study and declining with each succeeding grade. Consequently, the same individuals averaged at the 25th NCE (12th percentile) by grade 6 (p. 138). Various bilingual models were shown to be the most effective over the long-term, with students in bilingual programs “outperforming those schooled monolingually” (p. 73). The study indicated that student achievement in reading was “clearly the highest in the two-way bilingual immersion schools” (p. 140). Even though students in mainstreamed content or sheltered programs scored higher initially upon exit, students who exited from transitional bilingual programs achieved “at a significantly higher level than students who received content ESL services” (p. 141).

A 2008 meta-analysis of 15 studies all conducted prior to 1980, supported the findings of the Thomas and Collier (2002) study, indicating a statistically significant advantage for bilingual programs over English-only programs in their effect on reading outcomes (August, et al., 2008). This same analysis found that oral English proficiency levels and opportunities for interaction and output correlated with higher reading outcomes (p. 163). The researchers highlighted a
study by Klingner and Vaughn (1996) that indicated that “children with the potential to benefit most from the [reading] intervention had some initial reading ability and fairly high levels of second-language oral proficiency” (In August, et al., 2008, p. 163).

A recent study by Butler and Hakuta (2009), investigated the relationship between fourth-grade ELL students’ academic oral proficiency and their academic reading proficiency. The study focused on students who already proved to be orally proficient in conversational English. Butler and Hakuta found that teachers in the mainstream setting tended to underestimate the reading and language needs of students who had well developed oral conversational skills. They suggested that students “might benefit from direct instruction inasmuch as it helps them pay attention to the syntactic forms and expressions that the students are expected to use in schools for discussing academic concepts” (p. 434). Their findings were consistent with the work of Long (1996) and Swain (1985), concluding that students had “incomplete knowledge of forms and limited productive skills” (p. 434) despite displaying basic oral competency. Even though immersion and a focus on providing comprehensible input may have some benefits for students’ receptive language, their productive language skills do not appear to be sufficiently addressed by being fully immersed.

A 2008 study in Washington State tracked state assessment scores for ELL students in a large school district for three consecutive years after the students attained Transitional (Exit) level 4 on the Washington Language Proficiency Test (WLPT) and were no longer eligible for ELL services. These students, on average, scored significantly higher than the district and state averages in reading and writing on the Washington Assessment of Student Learning (WASL) (De Leeuw, 2008).
More recently, on the Spring 2010 state wide assessment, 53 percent of ELLs who transitioned out of ELL support services based on their 2010 WLPT-II scores also met the reading standard on the state-wide assessment for all students. Additionally, 63 percent of the transitioned ELLs also met the writing standard in that same year (Malagon, McCold, & Hernandez, 2011a, p. 28). These results would indicate (not surprisingly) that as students gain English proficiency, they are more likely to achieve at comparable levels to their native English-speaking peers on high stakes assessments in literacy. It is important to keep in mind that language acquisition is a long process, taking as long as 10 years to gain fluency, especially if the student’s learning environment is not optimal or if literacy in their first language is not supported.

**Mainstream v. Sheltered Instructional Settings**

A 3 ½ year ethnographic study by Harklau (1994) sought to identify how second language learners fare as they transition from the sheltered instructional setting into the mainstream classroom. Harklau observed four Chinese immigrant students in one San Francisco high school as they participated in both the ESL classroom and the mainstream classrooms. The study contrasted differences in spoken and written language use, content and curricular goals, instructional methods, feedback, and socialization factors, challenging the “folk belief that children will learn English faster if they are in regular classes with native speakers of English” (p. 242).

Harklau (1994) documented “… only 8 instances of learners talking in (mainstream) class discussions over the course of 12 days of classes, and 10 dyadic exchanges with teachers” (p. 251). Harklau also found that students were less engaged in the mainstream classrooms and were afforded fewer opportunities for comprehensible input and comprehensible output. Since
the need for comprehensible input and opportunities to engage in comprehensible output are widely accepted factors for language acquisition, mainstreaming students too early may in fact hinder their rate of acquisition.

Another study conducted by Langman (2003) investigated the effects of a model change that moved middle school ELL students from a sheltered ESL program to a model that fully mainstreamed ELLs into the regular classroom. Langman found that, even after all teachers in the school had received ESL training and incorporated ESL strategies into their teaching to make input (and content) comprehensible, students’ academic language did not improve. The study also indicated that because students did not receive specific English language instruction, they showed little improvement in their English abilities over the preceding year. Langman concluded that the lack of focused language instruction resulted in only “incidental language learning opportunities” (p. 20).

Language Anxiety in the K-12 Classroom

Based on the widely accepted premise that high levels of anxiety will interfere with an individual’s ability to acquire language, Pappamihiel (2002) compared English language anxiety levels for 178 sixth through eighth grade ESL students in two different instructional settings. The hypothesis was that anxiety levels would be higher when students were in the mainstream classrooms as compared to the ESL classroom. The participants were 178 Mexican immigrant students (91 female/87 male) across 7 different schools in a major city in Texas. All of the participants had been in the U.S. at least one year and were in mainstream classes for part of the school day.

To determine levels of anxiety in each of the two settings, the English Language Anxiety Scale (ELAS) was administered, followed by homogenous focus group discussions by the
participants. Pappamihiel (2002) correlated anxiety levels with various categorical and academic achievement factors. The results showed that there were significant differences between the ELAS scores in the ESL and the mainstream classrooms, with overall anxiety being much higher in the mainstream classes. Given the widely accepted negative correlation between anxiety and language learning, this study contributes greatly to the mainstreaming v. sheltered debate, giving credence to sheltered instructional models over full mainstreaming.

**Teacher Perceptions and Beliefs**

An integral part of any learning environment is the teacher. As we all know, the beliefs, attitudes, and perceptions of the classroom teacher can significantly impact the learning taking place in that environment. Teacher attitudes toward non-English speaking students can likewise impact the level of interaction in the mainstream monolingual English classroom. In a study in the United Kingdom, Franson (1999) interviewed three classroom teachers of primary grade level students from three different schools in the London suburban area to ascertain what perceptions and beliefs classroom teachers had about the language learners they taught. Although a small study in England, Franson’s findings are consistent with other studies on teacher perceptions conducted in the United States. Issues of underfunding, lack of training, lack of appropriate materials for English language learners, and lack of sufficient planning time and collaboration with English language specialists are concerns not solely limited to teachers in American classrooms.

Franson (1999) found strong concerns by the teachers pertaining to the well being of the EAL (English as an Additional Language) students in the mainstream classroom. Teachers in the study indicated that:
• “Sometimes children can feel embarrassed for example if they don’t know certain vocabulary…if you could do that privately on a one-to-one it might be better.” (p. 65)

• “A lot of the benefits (of mainstreaming) are for the mainstream children who are not English as a second language.” (p. 65)

Even though it was clear that the teachers in Franson’s study cared about their students, the teachers expressed overall feelings of resentment, resistance, apprehension, and fear, as well as being “overwhelmed.” Based on these findings, Franson concludes that “despite the well intentioned rhetoric, one might argue that language and learning needs of EAL (English as an Additional Language) pupils are not always well served by ‘mainstreaming’” (p. 59), ending with the suggestion that “mainstreaming EAL pupils may have granted EAL pupils equality of presence, but has not necessarily secured equality of participation” (p. 70). At least two questions remain: Do the benefits of mainstreaming students early outweigh the drawbacks?

**History of Mixed Methods Design**

Although it may seem that mixed methods research is something new on the scene, it is really just a case of something old becoming new again. Even Galileo, considered the father of the scientific method and experimental research, no doubt started out by taking anecdotal notes and observing the world in a very qualitative manner. While he may not have employed a complex coding scheme as delineated by Boeije (2010) or Freeman (1998), he most certainly looked for salient patterns before developing ways to test, measure, and verify what his qualitative observations would seem to indicate. In that sense, one might expect that he would
have employed a sequential exploratory design, **QUAL \(\rightarrow\) quan**, similar to one of the designs described by Creswell (2009).

**From Pragmatism to Paradigm Wars**

Moving ahead 300 years, to the late 1800s, we come to the age of classic ‘*pragmatism’*, a term introduced into philosophy by Charles Sanders Peirce in 1878 (Barnhart, 1995). Two other key figures from this early period were William James and John Dewey (Hookway, 2010). As the name certainly indicates, ‘*pragmatism’* has at its core something active and useful. In the words of Peirce (1879), pragmatism requires that we “consider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have. Then our conception of these effects is the whole of our conception of the object” (in Johnson & Onwuegbuzie, 2004, p. 17). (It’s rather interesting and a bit ironic to note, as Johnson and Onwuegbuzie do, that this quote is found in a work by Peirce entitled *How to Make Our Ideas Clear*.) Dewey (1920) clarifies, stating that “in order to discover the meaning of the idea [we must] ask for its consequences” (in Johnson & Onwuegbuzie, 2004, p. 17). In short, everything has some practical consequence, therefore, we need to examine these practical consequences in order to get at the truth.

As objectivism became the dominant research paradigm of the 20th century (prompted no doubt by the rapid advances in the sciences and in technology), it became generally accepted that all things could be quantitatively and objectively measured. It was assumed that reality was reality and that things like culture and environment should be not be part of the formula. According to Ayn Rand (1962), “Reality exists as an objective absolute—facts are facts, independent of man’s feelings, wishes, hopes or fears” (n.p).
During the 1960s and 1970s, qualitative research began to make more of an appearance, owing much to cultural factors of the period such as equal rights, woman’s rights, increased diversity, and bilingual education. In 1967, as qualitative research began to unfold and move beyond descriptive ethnography, Glaser and Strauss published their book, *The Discovery of Grounded Theory: Strategies for Qualitative Research* (Boeije, 2010). This would signal a move toward legitimation of qualitative research.

Often reflecting a constructivist or interpretivist paradigm, qualitative research, which consisted primarily of ethnography at the time, would be seem to come into direct conflict with objectivist views like those expressed by Rand (1962). According to Boeje (2010), Glaser and Strauss saw social science research as being “dominated by hypothesis testing and…devoid of any connection to everyday reality” (p. 8). Ethnographic research, at the other extreme, was “too preoccupied with description instead of explanation of social phenomena” (p. 8). This step toward systematizing qualitative research analysis served to help legitimize the field while simultaneously making it more difficult to put aside. The result would be the ‘Paradigm Wars’, which would continue into the new millennium.

Mixed methods research was beginning to gain a foothold in research practice during the 1990s and the beginning of the 21st century as evidenced by the increased number of mixed methods textbooks being published (Hanson, Creswell, Plano Clark, Petska, & Creswell, 2005). Yet, the debate over the validity of qualitative vs. quantitative research was continuing to roar in the background. In 1989, Gage stated that these “‘Paradigm Wars’ had come to a sanguinary climax” (p. 4). Unfortunately, 1989 in no way marked the height of the epistemological dispute. Both sides continued to argue vehemently for the superiority of their respective stances and their accompanying research designs (Anderson & Herr, 1999; Barone, 2001; Gage, 1989; Mayer,
2000, 2001). This was most particularly the case in the arena of educational research. The release of a report by the National Reading Panel (NRP) in 2000 indicated:

The evidence-based methodological standards adopted by the Panel are essentially those normally used in research studies of the efficacy of interventions in psychological and medical research. These include behaviorally based interventions, medications or medical procedures proposed for use in the fostering of robust health and psychological development and the prevention or treatment of disease. (NRP, 2000, p. 1)

The privileging of quantitative studies in the National Reading Panel report, with qualitative studies specifically excluded, prompted one of the members of the panel, Dr. Joanne Yatvin, the committee’s only teacher practitioner, to write a letter to the panel along with a 5-page statement of Minority View. In her statement, Dr. Yatvin points out several exclusions of the report, the very beginning of which is the criteria for what qualified as valid ‘research’ and which studies would be included. The specific exclusion of any ethnographic studies, with the inclusion of only empirical, experimental studies would set the definition for ‘research’ in the No Child Left Behind legislation the following year. She also pointed out the exclusion of anything in the report pertaining to research on the effects of home culture on children’s literacy development. Dr. Yatvin sums up the Minority View with the following statement:

In the end, the work of the NRP is not of poor quality; it is just unbalanced and, to some extent, irrelevant. But because of these deficiencies, bad things will happen. Summaries of, and sound bites about, the Panel’s findings will be used to make policy decisions at the national, state, and local levels. Topics that were never investigated will be misconstrued as failed practices. Unanswered questions will be assumed to have been answered negatively. (National Reading Panel (U.S.) , 2000, p. 2 of Minority View)
The information contained in the Minority View was not included at all in the 30-page Executive Summary of this document, which is likely what most legislators saw and used to develop legislation. Most legislators were, therefore probably unaware of the strong bias contained in the report.

The subsequent 2001 reauthorization of the Elementary and Secondary Schools Act requires that ‘scientifically based-research’ under No Child Left Behind [Section 9101(37)] must be “evaluated using experimental or quasi-experimental designs … with a preference for random-assignment experiments or other designs to the extent that those designs contain within-condition or across-condition controls” (iv). Such narrow definitions for research coupled with new levels of accountability fueled the academic debate even further. The November 2002 issue of *Educational Researcher* ran a rather heated series of articles and responses on the topic of what qualifies as valid research in education (Berliner, 2002; Erickson & Gutierrez, 2002; Feuer & Towne, 2002; Feuer, Towne, & Shavelson, 2002; Pellegrino & Goldman, 2002). The ‘Paradigm Wars’ appeared to be heating up even further. A similar debate appeared less than 5 years ago in the August 2006 issue of *Adult Education Quarterly* (Feuer, 2006; St. Pierre, 2006).

Today, in 2012, things may be coming together a bit. There is a strong contingent of researchers in the field of educational research (Hanson, Creswell, J.W., Plano Clark, Petska, & Creswell, J.D., 2005; Tashakkori & Teddlie, 2003) who agree with Johnson & Onwuegbuzie (2004) in the contention that mixed methods research situated in a pragmatic paradigm “offers a practical and outcome-oriented method of inquiry that is based on action and leads, iteratively, to further action and the elimination of doubt” (2004, p. 17). Despite the academic debate, “practicing researchers frequently ignore what is written by methodologists when they feel a mixed approach will best help them to answer their research questions” (p. 22). There is a
recognition (a renewed recognition perhaps) that research should follow what ever method or combination of methods is best suited to answering the research question.

It is quite ironic that more than 100 years prior to the current debate, William James (1907) published a series of lectures entitled “Pragmatism: A New Name for an Old way of Thinking” (in Hookway, 2010). In the first lecture of the series, he identified ‘The Present Dilemma in Philosophy’ as being, according to Hookway, a “fundamental and apparently irresoluble clash between two ways of thinking about things” (n.p.), namely the ‘tough minded’ and ‘tender minded’. Just as mixed methods research promises today to reconcile the ‘paradigm wars’ that have pitted quantitative research against qualitative research over much of the past 25 years, James (1907) argued that pragmatism was the way to overcome the same ‘dilemma’ a century ago. It would seem that everything old is new again and again.

**In Practice**

Along with the understanding that the grounding of mixed methods research in the pragmatic paradigm is not a new phenomenon, we also need to recognize that it is not new in practice either. Researchers particularly in the social sciences have long used multiple methods of data collection, analysis, and triangulation to understand a problem. A 4-year study by Thomas and Znaniecki, *The Polish Peasant in Europe and America*, was a mixed methods study conducted from 1918 to 1922 (Brannen, 2009). Other early examples of studies that contributed to bringing mixed methods into the current discussion include Campbell and Fiske’s 1959 study on the validation of psychological traits, which introduced a multitrait-multimethod matrix to the field. Later work by Jick, in 1979, brought in triangulation of different quantitative and qualitative data sources (Creswell, 2009). Another work published by Cook & Reichardt during that same year identified ten ways to combine qualitative and quantitative data, thus ushering in a
period of increased standardization while the epistemologically heated battles of the paradigm wars simultaneously ensued.

As mixed methods research continues to be increasingly part of the research landscape, resources and references to it continue increase as well. In 2005, Hanson et al. identified ten textbooks devoted to mixed methods research, three online or print journals, and one dedicated website. In the years since, many more have come about. A Google search of the terms “mixed methods research” and “integrating research,” for example, yields more than 288 thousand and 1.82 million hits respectively (Google, 2011). Google Books shows 30,500 results for books that include “mixed methods” across a variety of subject areas (e.g. nursing, business, biology). Online discussions number more than 3,300 and even include the area of engineering. Interest in mixed methods research, whether viewed as a coming together of two opposing paradigms or as third stand-alone, is apparently alive and well.

**In the Field of English Language Learning**

As research in general has moved from a more quantitative paradigm toward a mixing of data collection, analysis, and triangulation through a mixed methods approach, trends in second language research seem less well defined. Some argue that qualitative research either is on the rise or already dominates, while others take the opposite stance telling us that “parametric statistical procedures still ‘reign supreme’ ” (Lazaraton, 2005). So, which is it and where is mixed methods? The answer would seem to be the refrain so often heard in the context of second language teaching: “It depends.”

For the purposes of this discussion, I begin with the supposition that much of the research in the area of English language learning has come from a predominantly qualitative orientation. It has been my experience, as I search for studies related to my topic, that qualitative studies
seem to abound, while definitively quantitative studies are much harder to come by. I will begin, therefore, with a definition of qualitative research as put forth by Boeije (2010):

The purpose of qualitative research is to describe and understand social phenomena in terms of the meaning people bring to them. The research questions are studied through flexible methods enabling contact with the people involved to an extent that is necessary to grasp what is going on in the field. The methods produce rich, descriptive data that need to be interpreted through the identification and coding of themes and categories leading to findings that can contribute to theoretical knowledge and practical use. (p. 11)

Hinkle’s (2005) *Handbook of Research in Second Language Teaching and Learning* identifies the discipline of applied linguistics, along with research in second language processes, as beginning relatively recently, in the 1950s. Hinkel specifically identifies five areas of research in second language teaching and learning. Four of these areas fall into what are typically considered *qualitative* forms of research. The fifth area identified is *quantitative* research. Mixed methods is, surprisingly, not included at all.

Hinkel (2005) identifies ethnographic/qualitative, case study, and quantitative methods in second language research as being based on already established research paradigms, while classroom research and action research are seen as specific adaptations of these other methods to the context of second language teaching. It should be noted that Harklau (2005), whose chapter on ethnographic research begins the handbook’s section on methods, differentiates between ethnographic and qualitative research, whereas Hinkel has apparently indicated them to be interchangeable. Harklau also places *case study* and *classroom research* as types of *ethnography* rather than as completely separate categories of research.
Ethnography

In the 2005 *Handbook of Research in Second Language Teaching and Learning*, Harklau presents a comprehensive background of ethnographic research as one of the major approaches to second language research over the past 30 years, strongly influenced by the early disciplines of cultural anthropology, social anthropology, linguistic anthropology, and sociolinguistics. As a research methodology, ethnography had its beginnings in the field of anthropology, being particularly dominant from the 1870s to the 1920s. As part of the emerging disciplines of cultural and linguistic anthropology, many early ethnographies were conducted among Native American communities and in ‘primitive’ cultures of the colonies of the British Empire.

Preceding such early ‘ethnologists’ was Henry Schoolcraft (1793-1864), who documented his life among the Lake Superior Chippewa from 1822 to 1847. It is an interesting aside that Henry Wadsworth Longfellow based his poem, “Song of Hiawatha,” on Schoolcraft’s work. Longfellow reportedly stated, “I pored over Mr. Schoolcraft’s writings nearly three years, before I resolved to appropriate something of them to my own use” (Massie, 1856/1984). Research can certainly have far-reaching and unexpected effects. Despite some of the researcher bias that appeared in Schoolcraft’s 1856 work, he emphasized the importance of context and connectivity in trying to understand various aspects of culture, including the use and construction of language as a reflection of culture.

Progressing into the early part of the 20th century, ethnographic work in anthropology and sociology began to become part of an established discipline of cultural and linguistic anthropology, contributing in large part to the development of ethnographic methods in language research today. Most notably, from this period, is Malinowski’s contribution of the importance of participant observation in data collection and the beginnings of a distinction between
description and analysis in ethnographic research (Bronislaw Malinowski--Wikipedia, 2011). The distinction between description and analysis is still a contention in language research today.

In the 1960s, significant contributions were made by Gumperz and Hymes, who initiated the concept of the ‘ethnography of speaking’ in their 1964 studies. Their work moved linguistic analysis from focusing only on linguistic code to analysis of the communicative event (Harklau, 2005). By situating the method in ethnography, speech acts become analyzed in the context of setting, participants, purposes, topics, and codes. Interactions began to be seen not as culturally isolated but taken as culturally situated. Ethnography of communication research emphasizes samples of audio- or video-recorded interactions which are then analyzed in detail on aspects of ‘turns at talk, speech acts, sequences of speech acts, interactional encounters, speech events, social occasions, speech situations, and other aspects of communicative practices’ (Keating, in Harklau, 2005, p. 182). This area of research that focuses on the ethnography of communicative events continues to have great influence in the study of language interaction today (Chapelle, 1998; Gass, 1997, Harklau, 1994; Long, 1996, Swain, 1985).

What differentiates this methodology from previous ethnographic approaches in cultural anthropology and sociology is the ability provided by new technology to use recording devices to collect data. One major advantage of this type of data collection is the ability to replay the recordings in order to offer detailed analysis, rather than relying strictly on field notes. A disadvantage might be that it could create distance from the culture in which the speech event occurs, thus invalidating it as a true ethnography. It is important, therefore, that any analysis of recordings should also be accompanied by extensive field notes, either in written or recorded form, by the observer.
Harklau (2005) provides a very extensive listing and description of studies in language learning that can be categorized as *ethnography* using a wide range of data collection techniques (e.g. interviews, focus groups, discourse analysis, surveys, questionnaires, etc.). As I have previously indicated, *case study* and *classroom research* are identified as subcategories in the broader notion of *ethnography*. She points out, quite rightly, that, regardless of what other data collection techniques are employed, to qualify as *ethnography*, it must be based on participant observation.

**Mixed Methods?**

It would seem to be unclear whether second language research is or is not moving toward the use of mixed methods. A study by Henning (1986) inventoried articles published in two top journals in the field, *TESOL Quarterly* and *Language Learning*, to determine what types of studies, quantitative or qualitative, were being published. In the midst of the ‘paradigm wars’, mixed method studies were not identified as a possible category. The articles were analyzed and tabulated from 1970, 1975, 1980, and 1985. The data for *TESOL Quarterly* indicated an increase in the percentage of quantitative articles published over that period. According to Henning, in 1970, 12% of the articles were quantitative, with 88% categorized as qualitative in nature. By 1985, it was found that the emphasis had shifted toward a higher percentage of quantitative articles, with 61% quantitative and 39% qualitative (p. 704). In *Language Learning*, Henning found the percentage of quantitative research contained in the journal to have increased from 24% in 1970 to nearly 100% in 1985. Henning viewed this trend toward the dominance of quantitative research “as a positive development—a kind of coming of age of a discipline” (p. 704).
In the same 1986 issue of *TESOL Quarterly*, Chaudron and Wolfson each suggest a different trend. Chaudron found it “encouraging that the variety of recent approaches to L2 classroom research has expanded” (p. 719), citing studies that used a variety of data collection methods including ethnographic observation and simulated classroom activities. Chaudron’s article, “The Interaction of Quantitative and Qualitative Approaches to Research: A View of the Second Language Classroom” advocated, as indicated by the use of “interaction” in the title, for the integration of research methods, pointing out, “these studies will be of little benefit if we fail to recognize the value of both qualitative and quantitative research” (p. 719). Wolfson was also “convinced that a two-pronged approach toward data collection and analysis is necessary” (p. 697). Neither of them label what they are seeing and advocating for as “mixed method”, they did however seem to indicate that movement toward mixed method was the direction second language research in 1986 was and should be headed.

Moving ahead twenty years, in 2005, Lazaraton conducted a study, similar to Henning’s 1986 study. Lazaraton analyzed and categorized 524 research articles published in four prominent language learning journals over an 11-year period, from 1991 to 2001. Her study left more questions than answers. While the data tables in the article clearly showed that 86% of the published articles (450) over the 11-year period were categorized as ‘Qualitative’, with 13% ‘Quantitative’, and only 1% categorized as ‘mixed’, the text of the article stated “450 (86%) were quantitative in nature, 67 (13%) were qualitative, whereas the remaining 7 (1%) were a mixture of both” (p. 214). The description of the data and the information in the tables contradict each other throughout the article, leaving the reader with the quandary of whether to believe the numbers (data tables) or the description. How ironic for a chapter specifically titled “Quantitative Research Methods.”
I found this study to have two additional concerns. One is that the author did not seem to have a clear set of criteria for categorizing articles as ‘mixed’, stating:

In some cases it was unclear how to categorize approximately seven articles; in these cases, the data were analyzed quantitatively, but the majority of the article consisted of quotes from learners, transcripts, etc. These articles were categorized as ‘mixed’. (p. 213)

Since mixed methods had been around for more than twenty years prior to the 2005 study, I would have expected the researcher to have considered this category *a priori*.

My second concern is with the researchers choice of which journals were included and which journals were explicitly excluded from the study. Although the author identified that she was looking at research in the field of applied linguistics, she specifically excluded the journal of *Applied Linguistics* from her analysis, reasoning, “in my experience, it is not widely available in U.S. universities” (p. 218). This would not seem to be a very compelling reason to leave out a major journal. Lazaraton did acknowledge that a broader range of research methodologies are being employed, but that they “are not being published in journals at all” (p. 219) or may be included in what she termed ‘niche’ journals.

So, is mixed method research waxing or waning in second language research? It depends on where you look and what criteria is being used to determine if a particular study is or is not considered to be ‘mixed method.’ The field of second language research is very large and, like our student population, also very diverse. Certain methodologies may be more prevalent in certain specialty areas (e.g. adult education, workplace English, English for specific purposes, sociolinguistics, k-12 education, bilingual education, intensive English, etc.). In k-12 ELL education, I see the trend moving toward mixed methods, as teacher research and action research
become more prominent (Freeman, 1998). In contrast, state and federal governments continue to push for more evidence-based research.

**Challenges of Mixed Methods Designs**

Creswell (2009), along with Johnson and Onwuegbuzie (2004), identify three primary challenges faced by the researcher in employing a mixed methods approach. First is the need for extensive data collection. Since both quantitative and qualitative data are needed, the collection of multiple data sets is necessary to carry out the mixed methods study. With this in mind, Johnson and Onwuegbuzie recommend that mixed methods research may require a research team rather than a singular researcher, especially in concurrent designs. The second challenge is that mixed methods research is a highly time-intensive approach when it comes to analyzing the data. Since the data is collected and documented in both text and numeric formats, ample time to thoroughly analyze both forms is necessary. Johnson and Onwuegbuzie further indicate that it can be a more expensive model as well because of the additional time requirements and the use of multiple researchers. The third challenge is for the researcher to be familiar with both quantitative and qualitative forms of research. While collaboration may provide a solution to this challenge, the lead researcher still needs to have sufficient background in both forms to determine how each serves to answer the overall research question and effectively bring them together into one study. Johnson and Onwuegbuzie (2004) make the point quite nicely:

Today’s research world is becoming increasingly interdisciplinary, complex, and dynamic; therefore, many researchers need to complement one method with another, and all researchers need a solid understanding of multiple methods used by other scholars to facilitate communication, to promote collaboration, and to provide superior research.

(p.15)
Johnson and Onwuegbuzie (2004) go on to identify several additional challenges not brought out by Creswell. They warn that critics of the research may “contend that one should always work within either a qualitative or a quantitative paradigm” (p. 21). Although this may be a valid concern, as mixed methods become more accepted, paradigmatic concerns should also subside. Johnson and Onwuegbuzie also indicate that certain details regarding mixed methods in general need to be addressed, such as how to qualitatively analyze quantitative data and how to interpret data that indicate conflicting results.

**The Current Study**

Despite the challenges presented, a mixed method design offers several advantages over either quantitative or qualitative designs alone in answering my research questions on ELL student interaction in the mainstream and sheltered instructional settings. By using a mixed method design, I’ve been able to take advantage of four benefits identified by Johnson and Onwuegbuzie (2004, p. 21):

1. Answer a broader and more complete range of research questions;
2. Provide stronger evidence for a conclusion through convergence and corroboration of findings;
3. Add insights and understanding that might be missed when only a single method is used; and
4. Produce more complete knowledge necessary to inform theory and practice

In order to obtain a clear picture of verbal interaction in each instructional environment, I employed a “concurrent embedded” design strategy as described by Creswell (2009). This design allowed me to collect “two types of data simultaneously” and “gain perspectives from different levels within the study” (p. 215). By bringing together data from program records,
structured field observations, and anecdotal notes, I have been able to gain a better understanding of the different instructional environments. Such understanding can be used to support recommendations to various stakeholders on which environment provides the best opportunities for student interaction, thus facilitating language acquisition and translating into student success in the classroom.
Chapter 3
Overview of Methods

“Everything that can be counted does not necessarily count; everything that counts cannot necessarily be counted.”

Albert Einstein

In order to investigate the frequency of English verbal interaction for beginning level English language learners in the upper elementary grades and to simultaneously document the behavioral and environmental factors that influence such interactions, I chose to conduct a mixed methods study in which I employed a concurrent embedded design as described by Cresswell (2009).

This methods chapter provides an overview of the participants, settings, and methods employed of data collection for study as a whole. Information specific to the quantitative and qualitative components of the study are included in subsequent chapters 4 and 5 respectively. This includes detailed information on quantitative and qualitative analysis of the different sets of data collected.

Contextualizing the Study: Instructional Setting as Independent Variable

Instructional contexts are vital to understanding key elements of language learning because learners often rely on the context to gain meaning for that which is unfamiliar. For this reason, the context or setting in which classroom learning takes place is of primary concern in this study. The instructional settings in this study function as the independent variable that acts upon the dependent variable of student interactions, aka utterances. The following descriptions will help situate the reader to better understand both the contexts and the variables of this study.
**The School District.** This study took place in a large school district in Washington State with a linguistically diverse ELL population. The ELL population of the district is constantly growing, with qualifying students representing 5 percent of the overall student population at the time of the study. Over the past two years, the district’s ELL population has been one of the two fastest growing in the state. It is also increasingly diverse with 60 different languages identified during the 2010-2011 school year. This represents an increase of 15% over the previous year (Malagon, McCold, & Hernandez, 2011b).

**The School.** The mainstream and pullout observations were conducted in a highly diverse low-income school with 90 percent of the students receiving free or reduced lunch. Students who qualified for ELL services represented 12 percent of the school population, a rate triple that of the district average (Washington State Report Card, 2010-2011). The school principal indicated that the school also had a highly transient student population, with a student turnover rate of over 50 percent from September to June of the previous school year. The school was in step one of school improvement, as defined by No Child Left Behind, at the time of the study.

Typical of the other elementary schools in this district, ELLs in this school receive ELL support services through a combination of the pullout and push-in models. From their first day, ELL students are placed in a mainstream English only classroom. Beginning level students typically meet with the certified ELL specialist two or three times per week for 30 to 45 minutes in groups of between two and five students. Once students have progressed to the intermediate level (Level 2), they are generally no longer pulled out for service with the ELL certified teacher. They then receive support through push-in services with a bilingual specialist or tutor who may or may not speak their native language.
School Site Selection. In order to increase the likelihood that the maximum number of participants could be observed in all three instructional settings, the selection of the school site for the mainstream classroom and pullout group observations was based on two primary factors: 1) an expected high rate of student participation in the summer language camp; and 2) a large existing population of beginning level English language learners. Language camp records from the preceding three years were reviewed to determine which schools would be likely sites. Two schools were found to have similarly high rates of participation. The selected school had both a higher rate of language camp participation and one of the largest beginning ELL populations in the district.

By selecting a school with a large ELL population, it would be possible to conduct all of the mainstream and pullout group observations in the same building, thus allowing better control over the unsystematic variation that would have been introduced had the students been in different buildings. Being in the same building, teachers had access to similar resources and trainings. Likewise, student participants received similar ELL support services, in both quantity and quality, provided to them by the same ELL teacher and bilingual specialist. Other confounding variables such as school climate and neighborhood influences were similarly consistent for all participants.

The Summer Language Camp. The sheltered ELL observations were conducted at a three-week summer language camp on the campus of a local private university. The camp was in its 13th year as a collaborative effort between the university and the school district at the time of the study. The program runs for 2 hours every afternoon, five days per week over three weeks in July. It is an optional program, open to English language learners in the school district from age 5 to adult. It is often the case that entire families come to the camp together. Students are
placed in an age/grade appropriate, multi-level classroom with other English language learners of varying language proficiency levels, from Non-English speaking (NES) beginners to fluent speakers. At the time of this study, the program served more than 240 English language learners, representing more than 24 different languages.

Instructionally, the camp follows a content-based instruction (CBI) model. The theme-based program emphasizes teaching English through content, typically science or social studies. The camp theme is often associated with a current issue or topic, which becomes the focus of the content throughout the camp. Since 2011 was established by the United Nations as the “International Year of Forests,” it was decided that the content theme for the 2011 camp would be “Forests.” This theme was then developed within each classroom based the grade level of the students.

The average class size of the observed classrooms was 34.4 students. Each certified ELL specialist worked with two graduate or undergraduate students from the university’s TESL certificate program to develop and teach lessons that were appropriate to the ages and language levels of the students in each multi-level classroom. What makes the camp unique for the elementary ELL students is that the sheltered setting of the camp immerses the English language learners in classrooms with other English language-learning peers. This is a very different scenario from their mainstream classrooms where they are often one of only a handful, if not the only, non-native English speaker in the room.

Participants

Seven English language learners, four male and three female, in the third through sixth grades participated in the study across three different instructional settings: ELL Pullout Group, Mainstream Classroom, and ELL Sheltered (Content-Based) Classroom. At the time that they
were observed in the pullout groups, in the spring of 2011, all participants were receiving English language support services at the beginning/high beginning level (Level 1) as determined by either their qualifying or annual score on the Washington Language Proficiency Test (WLPT-II) for that school year. One of the participants had achieved intermediate status (Level 2) by the time she was observed in the mainstream classroom in the fall of the following school year.

The seven participants’ home languages represented 5 of the 60 languages present in the school district at the time of the study. The languages of the participants included KaRen, Swahili, Arabic, Karenni, and Nepali. All of the participants were classified as refugees. Six participants were considered to have limited or no formal schooling (LFS) prior to arriving in the United States.

**Participant Selection.** Students were identified for the study based on a list, provided by the school’s ELL Specialist, of all 3rd – 6th grade Level 1 ELLs in the selected school site. The list identified 21 Level 1 ELLs in nine different mainstream classrooms. As illustrated in Figure 5, a total of 20 English language learners were observed between the three different instructional settings to arrive at the seven learners who had participated in all three settings. It is this group of seven participants who are the focus of this analysis.
Table 2 shows the participant breakdown by instructional setting. The analysis of the data focused on the interaction rates of the seven students who were observed in all three settings. Access to mainstream classrooms for observation, combined with whether students chose to attend the language camp, ultimately determined which students would be included in the study.
Table 2

All Participants by Instructional Setting, N=20

<table>
<thead>
<tr>
<th>Instructional Setting</th>
<th>Number of Participants</th>
<th>Gender</th>
<th>Number of Home Languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pullout Groups</td>
<td>17</td>
<td>7 Male</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 Female</td>
<td></td>
</tr>
<tr>
<td>Sheltered Language Camp</td>
<td>8</td>
<td>4 Male</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 Female</td>
<td></td>
</tr>
<tr>
<td>Mainstream Classroom</td>
<td>13</td>
<td>6 Male</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7 Female</td>
<td></td>
</tr>
<tr>
<td>All 3 Settings</td>
<td>n = 7</td>
<td>4 male</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 Female</td>
<td></td>
</tr>
</tbody>
</table>

Instruments and Materials

Each observation session was documented directly onto a pre-printed observation sheet. Using tick marks for each occurrence, interactions (utterances) were classified into the categories of student-to-student or teacher-to-student. A separate page on the observation sheet allowed for anecdotal notes by the observer. Observers noted the academic subjects being taught, the activities taking place, and any unusual or interesting circumstances that occurred during the observation that might influence participant interactions.

Small digital video cameras were used to record each observation. The video cameras were placed on small tabletop tripods while the observer tracked interactions through direct observation in the classroom. The video record was for the purpose of triangulating the data. To ensure inter-rater reliability, one of every three recorded observations was reviewed by a second observer. The recordings were also used to review any observed irregularities during the sessions. These were reviewed by the researcher and the assistant observers as needed after each observation session.
Recruiting and Training Assistant Observers

Assistant observers were recruited from the graduate TESL program at the university where the summer program took place. Two assistant observers volunteered to assist with mainstream classroom observations during April. Given the limited 10-day window available for observations during the summer program, four assistant observers were recruited to assist with data collection in the sheltered classrooms during July.

After an initial training session to introduce them to the study, the original two assistant observers also took part in trial observations conducted in the university’s intensive English program in order to test the cameras, make adjustments to the observation sheets, and help ensure inter-rater reliability for the actual study. It was found that the video cameras timed-out and shut down after 29 minutes and that columns on the observation sheet needed to be reformatted. Needed adjustments to the observation sheets were made and new video cameras were obtained prior to the first observations at the school site.

Prior to their first observation session, I met with each assistant observer to introduce them to the observation materials and equipment and to provide training on documenting interactions and confidentiality requirements. Observers were instructed on what defined an interaction and how to document the observation. To obtain data that captured typical student interaction and behavior, observers were instructed to be in the background as much as possible and to not follow participants outside of the regular classroom. Although all students in the class would be aware of the observer and the camera, the observers were instructed not to let the students know exactly who or what was being observed. With the assistance of first language interpreters, the participants were debriefed after all observations for the study were completed.
To ensure inter-rater reliability, I met with each new observer for approximately 45 minutes immediately following their first observation session to review the video and clarify any questions. Subsequent observations also include a time for review; these generally took much less time than the initial review session. Cameras and completed observation sheets were collected after each observation session to compile data and review the video.

An observer protocol sheet was developed and provided to each assistant observer during the initial training (See Appendix B.). This helped to clarify procedures and reiterate the importance of maintaining participant confidentiality regarding observation sheets and video.

**Approvals and Consent**

School district and Institutional Review Board approvals were obtained with waivers for prior consent regarding classroom observations. Since students were unaware that they specifically were being observed, their interactions were assumed to represent a cross-section of their typical behavior in that instructional setting. Debriefing of student participants, with the assistance of bilingual interpreters, took place after all observations for the study were completed. Information for teachers was also limited until the observations were completed. Prior to the study, each teacher was made aware that ELL student behaviors were being observed, but specific detail was not included until after the observations were complete. Complete information was disclosed at the time of the follow-up interviews with each teacher.

**Data Collection**

All observations took place between April and December of 2011. Each observer utilized the same observation sheet format as previously described, with each session also recorded on video for review. In all, 101 observations were documented for a total of 8085 minutes.
Appendix C includes an overview of the observations of all 20 students to arrive at the data needed for the seven who participated in all three settings.

Since it was not possible to conduct all of the mainstream observations during the spring due to the scheduling of mandatory state assessments, I controlled for the maturation effect that might occur with students’ acquisition of language over time by counterbalanced the sequencing of the observation settings. This was accomplished by observing in half of the mainstream classrooms during the spring and the other half of the mainstream classrooms during the fall (See Figure 6.). Because some of the mainstream observations occurred prior to the sheltered observations and some occurred after, any difference in mean interaction rate that might be attributed to language development was, therefore, controlled for in the overall data. Pullout groups were all observed over three consecutive days during the first week of the study in April. Sheltered classrooms were observed over a ten-day period in July.

Figure 6. Sequence of Observations

No observations took place during the first two days of the summer language camp to allow for adjustments in classroom numbers and allow teachers to establish routines prior to any observations taking place. Mainstream observations in the fall did not begin until October to allow for ELL qualifying tests to be administered and schedules to be established.

In addition to controlling for maturation within the group by counterbalancing the mainstream observations, the mainstream observation sessions were scheduled to include at least
one morning and one afternoon session for each participant. This allowed observations to occur over various subjects throughout the school day so as not to privilege certain subject areas or activities over others. By including both morning and afternoon observations of each student, it was also possible to control for varying activity levels during different parts of the day due either to fatigue or heightened activity. Observations of the pullout groups occurred throughout the day and included both morning and afternoon sessions. Because the sheltered language camp classes only occurred during the afternoon, it was not possible to adjust the times for these observations. In all cases, interactions were not documented during recess or time with specialist, such as music, physical education, or art.
Chapter 4  
Quantifying Interaction Frequency

The quantitative component of this mixed methods study uses a within-subject design to systematically investigate the frequency of English verbal interactions of beginning level 3rd-6th grade English language learners across three different instructional settings: the mainstream grade-level classrooms, ELL pullout groups, and the ELL sheltered instruction classrooms. The following research questions are addressed and hypotheses proposed regarding the frequency of English verbal interactions:

1. What are the differences in frequency of overall English verbal interactions between the mainstream, sheltered, and pullout settings?
   - **Hypothesis 1a:** The frequency of English verbal interactions will be lowest in the mainstream classrooms.
   - **Hypothesis 1b:** The frequency of English verbal interactions will be highest in the pullout groups.

2. How does the frequency of teacher-to-student English verbal interaction compare between the three instructional settings?
   - **Hypothesis 2a:** Given the low student to teacher ratio, the frequency of teacher-to-student English verbal interaction will be highest in the pullout setting.
   - **Hypothesis 2b:** Given comparable class sizes, there will be no difference in the frequency of teacher-to-student interaction between the mainstream and the sheltered ELL classrooms.
3. How does the frequency of student-to-student (peer) interactions compare between the three instructional settings?

- **Hypothesis 3a**: The frequency of peer interactions will be lowest in the mainstream classrooms.
- **Hypothesis 3b**: There will be no difference in the frequency of peer interaction between the ELL pullout groups and the sheltered ELL classrooms.

4. Are there significant differences between teacher-to-student and student-to-student interaction frequency within each setting?

- **Null Hypothesis \( H_0 \)**: There will be no difference in the frequency of student-to-student and teacher-to-student verbal interactions within each of the three settings.
- **Hypothesis 4a**: The ratio of student-to-student and teacher-to-student interactions will be highest in the sheltered instructional setting.

**Quantifying Interactions: What counts?**

Quantifying interactions can be a tricky thing. Yet, for the purposes of this study, it is essential. The term *interaction* generally conveys a sense of working together and an exchange of ideas (what every language teacher hopes for), whereas an *utterance* need not involve another individual nor an exchange of ideas. As a language teacher, I do, of course, have a preference for the use of the term *interaction*, but I find this does not clearly convey what we are actually looking for in this particular study. Therefore, following Crossley, Salsbury, and McNamara’s (2010) study on lexical cohesive devices in which they examine “the frequency of NNS utterances that trigger negotiations for meaning” (p. 59), the terms utterance and interaction will be used interchangeably to mean a bounded unit of speech. This unit can be a single word, a
sentence, or a paragraph in length. It is bounded by either silence on either side of the unit or by another individual’s utterance either preceding or following. In terms of a dialogue between individuals, this might be thought of as a “conversational turn” as is the case in the Crossley, Salsbury, and McNamara study.

**Analysis and Results**

Observation times varied from one observation session to the other and between the three different instructional settings. To control for variation in observation times, it was therefore necessary to normalize the interaction data in order to compare across the different environments. For this purpose, an interaction rate was calculated for each observation. Interaction rates were calculated as interactions per minute for each participant. This rate is used as the unit of comparison throughout the analysis.

Preliminary analyses were carried out using the tools in Excel to calculate interaction rates for each participant observation. A composite interaction rate was then calculated for each participant in each of the three instructional settings. More detailed analyses were carried out using SPSS software. The dependent or outcome variable in this study is interaction rate. The independent or predictor variable is instructional setting, which consists of three different levels: mainstream classrooms, ELL sheltered classrooms, and ELL pullout groups.

To test the assumption of normal distribution in each sample, a Kolmogorov-Smirnov test was run. Results indicated that the distribution of interaction rates in the mainstream classrooms, D(7) = .200, p < .05, sheltered classrooms, D(7) = .200, p < .05, and pullout groups, D(7) = .081, p < .05, did not violate the assumption of normality.
Differences in Overall Interaction Frequency

Given that the same participants (n=7) were observed in all three settings, a one-way repeated measure ANOVA was carried out to determine if there were significant differences in interaction rates between the three instructional settings. This was followed by dependent (paired samples) t-tests to determine specific differences in interaction rates between each of the settings.

One-way Repeated Measures ANOVA. Mauchley’s test indicated that the assumption of sphericity had been violated, $\chi^2 (2) = 7.07, p < .05$, therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .57$). The ANOVA results showed that student interaction rates were significantly affected by the instructional setting, $F (1.14, 6.83) = 16.61, p = .004$, $\omega^2 = .91$. Additional multivariate tests confirmed the significance of the results, with Pillai’s Trace indicating $V = 0.90, F (2,5) = 22.77, p = .003$ (See Table 3.).

Effect size was calculated using the following equation as indicated by Field (2009, p. 480):
Table 3

**Multivariate Tests**\(^a\)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Value</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
<th>Noncent. Parameter</th>
<th>Observed Power(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pillai's Trace</td>
<td>.901</td>
<td>22.774</td>
<td>2.000</td>
<td>5.000</td>
<td>.003</td>
<td>.901</td>
<td>45.549</td>
<td>.991</td>
</tr>
<tr>
<td>Wilks' Lambda</td>
<td>.099</td>
<td>22.774</td>
<td>2.000</td>
<td>5.000</td>
<td>.003</td>
<td>.901</td>
<td>45.549</td>
<td>.991</td>
</tr>
<tr>
<td>Hotelling's Trace</td>
<td>9.110</td>
<td>22.774</td>
<td>2.000</td>
<td>5.000</td>
<td>.003</td>
<td>.901</td>
<td>45.549</td>
<td>.991</td>
</tr>
<tr>
<td>Roy's Largest Root</td>
<td>9.110</td>
<td>22.774</td>
<td>2.000</td>
<td>5.000</td>
<td>.003</td>
<td>.901</td>
<td>45.549</td>
<td>.991</td>
</tr>
</tbody>
</table>

\(^a\) Design: Intercept  \(^b\) Exact statistic  \(^c\) Computed using alpha = .05  Within Subjects Design: Setting

**Dependent T-tests.** Results of the dependent (paired samples) t-tests, as seen in Table 4, indicated a significant difference in mean interaction rates between each of the instructional settings. One-tailed significance is reported for each as specific alternative hypotheses were being tested. As predicted, the mean interaction rate was significantly lower in the mainstream classrooms (M = .265, SE = .073) than in both the sheltered classrooms (M = .637, SE = .191) and the pullout groups (M = 2.031, SE = .367), with \(t (6) = -2.570, p = .021, r = .724\), and \(t (6) = -5.211, p = .001, r = .905\), respectively. When comparing the ELL pullout groups to the ELL sheltered classrooms, interaction rates in the pullout groups were found to be significantly higher than in the sheltered classrooms, \(t (6) = 3.311, p = .008, r = .804\). In each case, the effect sizes, which take into account sample size, calculated using \(r = [((t^2) / (t^2 + df))^{1/2}\), indicated very large effects. With effect sizes ranging from .724 to .905, well above the large effect threshold of .5 identified by Field (2009), it is clearly shown, in each pairing, that the independent variable of instructional setting had a highly substantial effect on the dependent or outcome variable of interaction rate.
Table 4

*Dependent (Paired Samples) T-test.* This table shows differences in overall mean interaction rates between each pair of instructional settings.

<table>
<thead>
<tr>
<th>Setting Pairs</th>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Std. Error Mean</td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>Pair 1</td>
<td>-.37257</td>
<td>.38359</td>
<td>.14499</td>
<td>-.72734</td>
<td>-.01781</td>
</tr>
<tr>
<td>Mainstream –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheltered</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 2</td>
<td>-1.76657</td>
<td>.89700</td>
<td>.33904</td>
<td>-2.59616</td>
<td>-.93698</td>
</tr>
<tr>
<td>Mainstream –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pullout</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 3</td>
<td>-1.39400</td>
<td>1.11397</td>
<td>.42104</td>
<td>-2.42425</td>
<td>-.36375</td>
</tr>
<tr>
<td>Sheltered –</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pullout</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The instructional environment clearly had a substantial impact on students’ verbal interaction rates. As students moved between the mainstream environment, dominated by native English speakers, and the environments of the ELL pullout group and the ELL sheltered classroom, the differences in frequency of English usage was profound (See Figure 7). The most notable difference was found between the ELL pullout group and the mainstream classroom, where individual students exhibited from 4 times to 167 times more frequent interactions in the ELL pullout than in the mainstream classroom. A further discussion of individual differences is provided in the discussion.
Teacher-to-Student and Student-to-Student Interactions

Now that overall differences in interaction rates between settings have been established, where within the overall interactions, do the differences occur? Are the differences in interaction rates due primarily to the students’ verbal interactions with teachers in each setting or are the differences primarily attributed to the ELL students’ interactions with peers? To answer these questions, the frequency data collected regarding teacher-student and student-student interaction was analyzed: first, to compare between instructional settings, then to compare within settings. Since the same students participated in each of the three instructional settings, paired samples t-tests were again employed to analyze the data.

Teacher-Student Interactions Between Settings. Results of the dependent (paired samples) t-test, as shown in Table 5, indicated no significant difference in the frequency of verbal teacher-to-student interactions between the mainstream (M = 0.152, SE = .045) and sheltered (M = 0.220, SE = .063) settings.
**Table 5**

*Dependent (Paired Samples) T-test for Teacher-Student Interaction Rates*

<table>
<thead>
<tr>
<th>Teacher – Student Interaction Settings</th>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Std. Error Mean</td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>Pair 1 Mainstream - Sheltered</td>
<td>-0.06814</td>
<td>0.11282</td>
<td>0.04264</td>
<td>-0.17248</td>
<td>0.03620</td>
</tr>
<tr>
<td>Pair 2 Mainstream - Pullout</td>
<td>-1.56514</td>
<td>0.66603</td>
<td>0.25174</td>
<td>-2.18112</td>
<td>-0.94917</td>
</tr>
<tr>
<td>Pair 3 Sheltered - Pullout</td>
<td>-1.49700</td>
<td>0.74444</td>
<td>0.28137</td>
<td>-2.18549</td>
<td>-0.80851</td>
</tr>
</tbody>
</table>

Even though the difference between the sheltered and mainstream settings was not found to be statistically significant, the mean interaction rate for ELL students with their teachers was still 45% higher when the students were in the sheltered ELL classrooms than when the same students were in the mainstream classrooms (See Figure 8.).
For the ELL pullout groups (M = 1.72, SE = .263), teacher-to-student interaction rates were significantly higher than both the mainstream (M = 0.152, SE = .045) and the sheltered (M = 0.220, SE = .063) settings, with $t(6) = 6.217, p < .001$, $r = .930$, and $t(6) = 5.320, p = .001$, $r = .908$, respectively. Very large effect sizes ($r$) again indicate a substantial impact of instructional setting on interactions. While class size did not appear to be a factor in comparing the mainstream and sheltered classrooms, both averaging between 22 and 35 students, the small group size of the pullout groups, averaging only 2.5 students per group, is likely to be a contributing factor to the significantly higher teacher-student interaction rate in this setting.

It should not, however, be misinterpreted that the small group size is the primary factor for this difference. A review of the anecdotal notes and video of the observations did not show any evidence of similar increases in teacher-student interaction during teacher led small groups in the mainstream setting. When students participated in small groups with the teacher in the

![Figure 8. Teacher-Student Interactions in Mainstream and Sheltered Settings](image)
mainstream classes, the interaction rates appeared to remain relatively low. An additional study focusing on small group interactions would need to be conducted to further investigate whether there were significant differences in small group interaction between the different instructional settings.

**Peer (Student-Student) Interactions Between Settings.** Whereas no significant difference was found in teacher-to-student interactions between the mainstream and ELL sheltered setting, Table 6 shows that there were, however, significant differences in peer (student-to-student) interaction rates between the mainstream and sheltered settings. Peer interactions in the mainstream (M = 0.114, SE = .042) were significantly less frequent than in the sheltered classroom (M = 0.400, SE = .129), t (6) = -2.504, p = .023, r = .715. The results indicate that ELLs interacted with their peers nearly four times more frequently in the ELL sheltered classrooms than in the mainstream classrooms. An even greater difference existed when the peer interaction rates of the mainstream were compared with the ELL pullout groups (M = 0.298, SE = .063), t (6) = -4.416, p = .002, r = .874. The data clearly shows peer interactions to be substantially less in the mainstream classroom than in either of the two ELL environments. The very large effect sizes of .715 and .874 for the ELL sheltered and pullout settings, respectively, indicate a substantial negative affect of the mainstream environment on peer interactions for the beginning level ELL students in this study.
Table 6

Dependant T-test for Peer (Student-Student) Interaction Rates Between Settings

<table>
<thead>
<tr>
<th>Student – Student Interaction Settings</th>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Std. Error Mean</td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>Pair 1 Mainstream – Sheltered</td>
<td>-.28671</td>
<td>.30292</td>
<td>.11449</td>
<td>-.56687</td>
<td>-.00656</td>
</tr>
<tr>
<td>Pair 2 Mainstream – Pullout</td>
<td>-.18443</td>
<td>.11050</td>
<td>.04176</td>
<td>-.28662</td>
<td>-.08224</td>
</tr>
<tr>
<td>Pair 3 Sheltered – Pullout</td>
<td>.10229</td>
<td>.30308</td>
<td>.11455</td>
<td>-.17802</td>
<td>.38259</td>
</tr>
</tbody>
</table>

When students were with their English language-learning peers in either the ELL sheltered or ELL pullout settings, no statistically significant difference was found in student-to-student interaction rates. Class size in the ELL sheltered classrooms was substantially larger, with an average of 34.4 students, whereas the average size of the pullout groups was only 2.5 students. Despite a substantial difference in group size, two important factors remained the same in these two ELL settings: 1) the students’ peers were also English language learners; and 2) the teachers in both the sheltered and pullout settings were specifically trained to work with English language learners. Although not shown to be significant, it is interesting to note that the difference in peer interaction rates favored the ELL sheltered classroom, with a 34% higher rate of verbal interaction despite the much smaller size of the pullout groups.

Figure 9 illustrates a comparison of peer interaction rates between the three settings. The mean peer interaction rate of the ELL sheltered classrooms alone is approximately equal to the peer interaction rates of the mainstream classrooms and the pullout groups combined. ELL
students clearly have much higher rates of English interaction with their peers in the ELL sheltered classroom than in either the pullout group or the mainstream.

**Figure 9.** Comparison of Peer Interaction Rates by Setting

### Differences Between Teacher and Peer Interactions Within Settings.

The final analysis considers the comparative frequency of teacher-to-student and student-to-student interactions within each of the three instructional settings. Paired samples tests were again used to discern differences in mean interaction rates (See Table 7.). The differences here are examined within each setting rather than between settings.

Results from the mainstream observations, with the lowest overall interaction rate of the three settings, indicated no significant difference between teacher-student (M = 0.152, SE = .045) and student-student (M = 0.114, SE = .042) interactions. Beginning ELL students interacted with their mainstream teachers only slightly more often than with their mainstream classmates. The low rate in the mainstream setting was not overtly influenced by neither teacher nor peer interactions. ELL students appeared to interact at low rates regardless of whether the interactions were with their teachers or their classmates in the mainstream.
Table 7

**Paired Samples Test for Teacher-Student vs. Student-Student Interactions Within Settings**

<table>
<thead>
<tr>
<th>Instructional Settings</th>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Std. Error</td>
<td>Mean</td>
<td>Lower</td>
</tr>
<tr>
<td>Pair 1</td>
<td>Mainstream T-S – Mainstream S-S</td>
<td>.03800</td>
<td>.12545</td>
<td>.04742</td>
<td>-.07802</td>
</tr>
<tr>
<td>Pair 2</td>
<td>Sheltered T-S – Sheltered S-S</td>
<td>-.18057</td>
<td>.22340</td>
<td>.08444</td>
<td>-.38718</td>
</tr>
<tr>
<td>Pair 3</td>
<td>Pullout T-S– Pullout S-S</td>
<td>1.41871</td>
<td>.71769</td>
<td>.27126</td>
<td>.75497</td>
</tr>
</tbody>
</table>

Results from within the two ELL settings (sheltered and pullout) showed significant, yet opposing, differences between peer and teacher-to-student interaction rates. In the ELL sheltered classrooms, student interactions with their peers (M = 0.400, SE = .129) were significantly more frequent than interactions with their teachers (M = 0.220, SE = .063), t (6) = 2.139, p = .038, r = .658. Peer interactions represented 65% of the verbal utterances in the ELL sheltered classroom. Conversely, in the ELL pullout setting, the teacher-to-student interaction rate (M = 1.717, SE = .263) was significantly higher than the student-to-student rate (M = 0.298, SE = .063), t (6) = 5.230, p = .001, r = .906. Peer interactions, in this case, represented only 15% of the total interactions in the pullout setting. These results indicate that the larger of the two ELL settings, the sheltered classroom, allowed for more frequent peer interaction; whereas, the small student-to-teacher ratio of the pullout groups allowed for a greater proportion of teacher-to-student interaction. Figure 10 illustrates differences within each of the three settings.
Figure 10. Teacher and Peer Interactions as a Percent of Overall Interactions
The Exceptional Case

Throughout this study, individual students showed substantially more verbal interaction in the sheltered ELL classroom than in the mainstream classroom. These differences ranged from twice to nine times the mainstream rate. There was, however, one notable exception. It is on this one exceptional case that I would like to expand upon here.

As can be seen in Figure 11, Poh Reh showed almost identical interaction rates in both the mainstream and sheltered classrooms. Not at all typical of the other ELLs in the study, his mainstream interaction rate was actually slightly higher in the mainstream. In the pullout group, with only a few ELL students and a highly trained, experienced ELL specialist, his interaction rate of 3.99 interactions per minute, was more than ten times what he exhibited in either the mainstream or sheltered classroom settings and more than double the rate of any other student observed throughout the study. A highly social fourth grader, full of energy and clearly very bright, his classroom teacher found him to be a joy in the classroom and expounded on his abilities in math. The ELL specialist described him as a “pleaser,” always seeking to make others happy.
During the second day of work in pairs to create visual representations on graph paper of multiplication factors. The generally talkative and jubilant beginning level ELL was paired with a confident native speaker (NS). The student teacher commented that he had paired the ELLs in the class with native English speakers to ensure that they had good models of English to rely on. The mentor teacher agreed stating that she “tries to make sure that the ELLs never work together.” Although this might at first seem to make sense, the results of such efforts to keep ELLs apart did not increase verbal interaction.

What I observed during this cooperative learning activity was that Poh Reh, who could hardly contain himself during non-instructional times in his four-person table group, spoke not one word during this mathematics activity with his native English speaking partner. Despite the
fact that he was highly competent in his mathematical abilities (according to his teacher), he contributed nothing to creating nor discussing the mathematical model with his partner. Instead, he allowed the native speaker to take over and do all of the talking and the work. This 30-minute period yielded no verbal interaction at all, thus no opportunity for the student to negotiate for meaning and subsequently reap the benefits afforded by such interaction.

In contrast, I noticed the interaction of a non-native speaking (NNS) pair of students working next to them during the activity. In this case, there were two intermediate level ELLs working together, one Marshallese and one Spanish speaker (a scenario, which the student teacher pointed out, was not his intention, but occurred because their English-speaking partners were both absent). With the two language learners working together in the NNS-NNS pairing, I observed that the students were highly engaged in the activity and freely shared their ideas with each other. With both ELL students asking questions and clarifying meanings, the negotiation cycle was fully engaged, with plenty of opportunity for equal participation. One might suggest that the increased interaction in this NNS-NNS pairing might be due to their intermediate level status. However, my own experience in the classroom, combined with other research (Mackey, 2012; Swain M., 2005; Varonis & Gass, 1985) on NNS interaction, attribute the difference to the non-native speakers comfort level in working with another English language learner, someone who might be considered their true ‘peer’.
Chapter 5

What’s Behind the Numbers? A Qualitative Analysis of Classroom Observations

April 2011

In a fourth grade classroom, on a rainy April day, Bluh sits alone at his desk. The previous day, three weeks prior to the state mandated assessments, the desks had been placed in rows to simulate the testing situation. They will remain as such for the next five weeks, until all the testing is completed. Other students are moving all around Bluh, seeking help from the teacher, working on projects, or reading together. As he sits alone, playing with the Pokémon toy in his desk, there is a whirlwind of activity spinning around him. It’s as if he is a statue in the midst of a hurricane, stoic and frozen at this desk, speaking to no one and no one speaking to him.

The preceding scenario occurred during the first day of observations in the mainstream classrooms and was the first piece of video that I viewed from the study. It is one of the scenes that stayed in my mind throughout the study and captures the essence of what would later be observed in other mainstream classrooms with other ELL students, reflecting and personalizing the results of the quantitative data discussed in the previous chapter.

After finding significantly low rates of English language interaction in the mainstream classrooms for the quantitative portion of this mixed methods study, I set out to look more closely at the individual behaviors in the mainstream and ELL sheltered instructional settings. Since the mainstream and sheltered classrooms both have large class sizes, averaging 22.5 and 34.4 respectively, it is upon these two instructional settings that I shift the qualitative focus of this investigation.
To provide further context for the English interactions and to better understand the impact of the two instructional environments, I proposed two qualitative research questions: 1) How did the students’ behavior differ between the mainstream and the sheltered instructional settings? and 2) What circumstances in the classrooms may have influenced the observed interactions?

Methods

Student Participants and Environment

The seven beginning-level English language learners observed in both the mainstream and sheltered classrooms all attended the same low-income elementary school in a large Washington State school district. There were four boys and three girls observed in the 4th–6th grade classrooms. Of the school district’s 60 different languages, these students’ home languages represented five: Nepali, Arabic, Karen, Swahili, and Kareni.

Despite being enrolled in age appropriate classrooms at the fourth, fifth, and sixth grades at the time of the classroom observations, six of the students were identified as limited formal schooling (LFS) prior to their arrival in the United States. This designation includes two students who had never attended school at all prior to being placed in the mainstream classroom of their present school. The students had all come into the United States as refugees within the previous two years. As the first observations of the study began, in the spring of 2011, the students had been in the American school system an average of 11.1 months, with some as few as two weeks and others as long as 22 months.

Table 8 provides information on the seven individual students who were observed in both the mainstream and sheltered classrooms. In each case, pseudonyms are used when discussing individuals in this study. The grade levels indicated on the table are the grades that the students
were in at the time that they were observed in their mainstream classrooms, with an “S” prefix indicating that they were observed during the spring and an “F” prefix indicating that the mainstream observation occurred during the following fall.

Table 8

*Individual Students*

<table>
<thead>
<tr>
<th>Student Name (Pseudonym)</th>
<th>M/F</th>
<th>Home Language</th>
<th>Grade Level</th>
<th>ELD Level</th>
<th>Prior Schooling</th>
<th>Months in US Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bluh</td>
<td>M</td>
<td>KaRen</td>
<td>S4</td>
<td>1</td>
<td>Limited</td>
<td>22</td>
</tr>
<tr>
<td>Christelle</td>
<td>F</td>
<td>Swahili</td>
<td>S6</td>
<td>1</td>
<td>Limited</td>
<td>10</td>
</tr>
<tr>
<td>Nadia</td>
<td>F</td>
<td>Arabic</td>
<td>S6</td>
<td>1</td>
<td>Yes</td>
<td>19</td>
</tr>
<tr>
<td>Hso</td>
<td>M</td>
<td>Karenni</td>
<td>S6</td>
<td>1</td>
<td>Limited</td>
<td>0.5</td>
</tr>
<tr>
<td>Sitara</td>
<td>F</td>
<td>Nepali</td>
<td>F5</td>
<td>1/2</td>
<td>Limited</td>
<td>11</td>
</tr>
<tr>
<td>Benjamin</td>
<td>M</td>
<td>Swahili</td>
<td>F6</td>
<td>1</td>
<td>None</td>
<td>10</td>
</tr>
<tr>
<td>Poh Reh</td>
<td>M</td>
<td>KaRen</td>
<td>F4</td>
<td>1</td>
<td>None</td>
<td>5</td>
</tr>
</tbody>
</table>

**Data Collection**

All sheltered observations occurred during July of 2011. Four of the students were observed in their mainstream classrooms during the spring, prior to observations in the sheltered classrooms of the summer program. The remaining three students were observed in their mainstream classrooms during the fall. By switching the order in which the mainstream and sheltered observations occurred for half of the students, I was able to control for any differences in student behaviors that might be attributed to the passage of time in the overall study, thus controlling for the confounding variable of maturation within the group.

Video of each observation was reviewed to ensure the reliability of the observation notes in capturing the atmosphere and behaviors in the classrooms. The video also served to provide examples and additional context for the observers’ notes. In two cases, the observers asked to review the video because they felt they may have missed something during the direct
observation. In each case, I viewed the video with the observer and discussed the context. It was determined that no changes to the observation data were needed.

Follow-up debriefing sessions were conducted with individual teachers and with the students once all of the observations were completed. I met with each teacher separately in order to be able to share information specific to their students and answer any questions they had pertaining to individual participants. For the students, we met in groups with the assistance of bilingual interpreters from the school district. The specific purpose of the study was shared during these debriefings along with some of the preliminary findings from the quantitative portion of the study. Both teachers and students were also given an opportunity to ask questions, discuss, and share their own feelings about the study. Some of the comments from these sessions are used here to enhance the understanding of the observation sessions.

Analysis

In order to gain deeper insight into the quantitative data described in the previous chapter, in which significant differences in the rates of English interaction were found between the mainstream classrooms and the sheltered ELL classrooms, the observation notes and reflective fieldnotes were systematically analyzed using a grounded analysis method. The use of grounded analysis with open coding was chosen, rather than axial or selective coding, in order to allow the themes in the data to emerge from the data itself. In this way, I was able to include coding and categories that may not have been anticipated in an a priori coding process. By grounding the codes in the data itself, I also sought to minimize my own bias in the analysis.

Based on the procedures for grounded analysis outlined by Freeman (1998) and simulating Boeije’s (2010) process of open coding, I analyzed the qualitative data using the following process:
1. Read through notes to gain an overall sense of the data.

2. Read through notes a second time, underlining and highlighting key words or phrases and adding working codes in the margins. A color-coded system was utilized to correspond to the codes derived from the data. This helped to make the code categories more salient once all of the notes had been coded.

3. Following the second reading, codes were collapsed or added to better reflect the emerging themes of the data. Requisite changes to the color-coding scheme occurred simultaneously.

4. Categories reflecting related codes were identified and placed in a table to be used to tabulate the frequencies of the occurrence of each code in the data and determine dominant themes.

5. Coded sections of text were placed into their appropriate categories for easy reference and retrieval.

6. Coded text were then used to create word clouds using the Wordle.net website.

There were several items in the notes that could have been coded and placed in two or more categories. In those cases when the sentence or phrase could be split without affecting the meaning of each section, this was done, resulting in two pieces of coded text. Frequently, however, this was not possible; in which case I coded the text and determined the category that best matched the phrase within the overall context of the passage. Thus, each coded section of text was only placed into one category. This same systematic process was followed to analyze the anecdotal data from the sheltered classrooms as well, with new codes added as necessary with the new data set.
**Word Clouds as Analytical Tools.** Although the use of word clouds, in this case Wordle, as an analytical tool for qualitative analysis has not yet entered the mainstream, I have chosen to use it here as a means not merely to display my findings, but to assist in pulling out salient themes in the observation notes of the mainstream and sheltered classrooms. According to McNaught and Lam (2010), the use of word clouds in qualitative research is particularly useful to conduct preliminary analyses of texts or, as I’ve used it here, as “a validation tool to further confirm findings” (p. 631). I am, therefore, employing Wordle as a means of triangulating the data obtained from the observation notes.

**Divergent Themes in the Findings**

The categories that emerged in the analysis of the anecdotal data followed completely divergent themes between the mainstream and the ELL sheltered classrooms. The observers’ notes from the sheltered classrooms painted a picture of happy, engaged ELL students eager to learn and practice their new language skills. Whereas the notes from the mainstream observations showed these same students to be largely disconnected from the instruction, often appearing unhappy, distressed, and frustrated.

**Mainstream Findings**

As seen in Figure 12, the notes from the mainstream classrooms revealed seven categories of observer comments: 1) Disengaged/Alone, 2) Assessment, 3) Singled Out, 4) Non-Academic, 5) Attempted Initiation, 6) Behavior Problems, and 7) Positives. Two particular categories stood out as the most dominant. These were the categories of Assessment and Disengaged/Alone, with 59% of the coded comments falling into these two categories.
Figure 12. Categories of Coded Data from Mainstream Observation Notes

**Disengaged or Alone.** The mainstream notes were overwhelmingly dominated by comments that indicated that the ELL students were disengaged or alone. Comments coded in this category represented 39% of the 126 pieces of coded text and included codes such as “head down”, “resting”, “angry”, “lost”, “alone”, and “hiding.” Any attempts to quantify these codes cannot adequately capture the overwhelming nature of this category and the impact of directly observing individual students undergoing these experiences. The scenario at the beginning of this chapter only begins to exemplify such cases. For each of the seven students, Bluh, Sitara, Christelle, Nadia, Hso, Benjamin, and Poh Reh, similar experiences of being disengaged or alone all were observed. This was, for nearly all of them, a regular experience during the mainstream observations. The following quotes from the observers’ notes illustrate the disengagement and
isolation of students as witnessed by the observers. Comments come from three different observers in seven different classrooms.

Bluh:

- 4/20 “The students then listened to the teacher read as they were to draw pictures. Participant did not draw a picture (played w/his watch and rested his head).”
- 4/21 “Students were specifically asked to turn to a neighbor and discuss… Subject’s neighbor did not have a partner, yet subject did not talk w/his neighbor or anyone else.

Christelle:

- 4/20 “Christelle has very little contact with any other students in the classroom. She is sitting next to someone, but doesn’t speak to her.”
- 4/22 “Christelle worked quietly at her desk, not communicating in any way with her group members.”

Nadia:

- 5/26 “Students moved to teams (GLAD) to work on maps, inquiry charts, timelines, etc. Nadia is working on the map alone.”

Hso:

- 5/31 “Hso is not engaged, playing with pencil and items on desk.”
- 6/1 “While students work on math activity, Hso is not focused, looking around the room and at pictures on the wall next to him.”

Poh Reh:

- 10/18 “Poh is playing with things in desk.”
- 10/19 “Poh is completely lost on what to do. Starts drawing on page.”
Benjamin:

- 11/16 “Ben moves himself away from the other students and puts himself next to the teacher and does not talk or interact with other students.”
- 11/16 “During team Brainquest, Ben is completely separate from the group.”
- 11/18 “Ben is sitting directly in front of the screen not near anyone.”

Sitara:

- 12/7 “Sitara is not listening to the story. She is shuffling through blank papers with her back turned to the teacher.”
- 12/7 “Sitara sits behind a chair on the floor not near anyone…Previously she sat behind the edge of a desk.”
- 12/7 “Sitara has very little peer interaction. The vast majority of her interactions are with the teacher.”

Although the mainstream teachers did not physically separate the English language learners from the other students in the classroom, the classroom observers documented that the ELL students often separated themselves either physically or mentally from the other students as well as from the instruction taking place.

Assessment. Although I scheduled the observations not to interfere with standardized testing, the second most dominant category for comments in the mainstream observations pertained to assessment. This included practice tests and formal assessments as required by the school district. Whether observations occurred in the spring or in the fall, in the morning or in the afternoon, very few of the observation sessions were without some sort of formal assessment or practice test. A native English speaking student in one of the classrooms commented on this overabundance of testing after the teacher announced yet another assessment, “How many of
these assessments do we have?” Another student chimed in, “All these assessments are getting tiring.”

For the beginning level English language learners, this extreme emphasis on assessment is particularly discouraging. According to State program guidelines and in accordance with federal No Child Left Behind legislation, qualified ELL students who have been in the United States school system for less than a year are exempted from taking the State’s annual reading and writing assessments (Malagon & Chacon, 2009, p. 15). The ELL students in this study who would qualify for such an exemption were not, however, observed to be exempted from taking school district required tests in these same areas. In addition, ELL students were not allowed any accommodations for the required district assessments even though such accommodations would be allowed on the annual State assessments.

This discrepancy did not go unnoticed by the classroom teacher, who shared with me her frustration that there was not better communication with the classroom teachers about which students would be exempted from the annual State assessments. She expressed further frustration that they were not better informed early in the year on what accommodations would be allowed for the ELL students. This timely information would allow her to incorporate accommodations into instruction throughout the year. Instead, she would learn this essential information only days before the assessments began. Similar frustrations were expressed by other classroom teachers throughout the study and during follow up debriefing sessions.

The Wordle in Figure 13 shows how dominant the word assessment was in the coded text of the mainstream observation notes. Its prevalence in the coded text illustrates how much assessment has also come to dominate the activities of the mainstream classroom. As a colleague recently reminded me, a basic premise of achievement testing is that we should test
only that which has been taught. For the beginning level ELL with little or no prior schooling, much of what they are being tested on in the mainstream classroom has not yet been taught.

![Wordle of Dominant Words in the Mainstream Observation Notes](image)

**Figure 13.** Wordle of Dominant Words in the Mainstream Observation Notes

**Singled Out.** Although this category may not have shown up frequently in the mainstream notes, the instances when it did were most disturbing. Again, as in the opening scenario, these examples involve Bluh, a level 1 KaRen refugee. The following are taken directly from the observer’s notes, including parenthetical comments:

Participant was called on to answer a question – he answered it correctly and a few students remarked, “Good job, Bluh!” (This was not repeated for any non-ELL students who answered questions – not with the same intensity at least).

Although it may appear that the students were being encouraging, which was likely their intent, it was clear from viewing the video that Bluh did not appear to feel “encouraged” by his classmates’ comments. He appeared, instead, to be very uncomfortable and distant. The sense one gets when viewing the video, and which the observer also shared with me later, as we reviewed the video, was a pervading atmosphere of condescension rather than encouragement.
The preceding example was immediately followed in the observer’s notes that same day by yet another incident. Throughout the week students had been working to earn points toward a Friday movie, when students asked if they would get a movie later that day, the observer noted the following:

Teacher stated that she was not going to show a movie because Bluh doesn’t understand it. Singled him out and made other students resent him (plus movies are a great way to learn English when used w/subtitles). Resentment evident by faces the students made.

One remarked, “That’s not fair!”

Upon viewing the video, I found that the teacher’s statement had also referenced the fact that the deaf student in the class “could not hear” as an additional reason not to show the movie. While the teacher’s efforts to “be fair” to the ELL student and the student with a hearing disability may appear somehow biased against these two students and ultimately “unfair” to the entire class, this instance clearly shows a lack of understanding by the teacher about what would benefit students in their acquisition of language whether their primary language is KaRen or American Sign. Despite having been in school 22 months (longer than any other student in the study), Bluh’s rate of interaction in this mainstream classroom was, not surprisingly, the lowest of all the students observed in the study.

**Non-Academic.** The mainstream category of “Non-Academic” includes comments by the observers that pertain to the ELL students using English during non-instructional times such as standing in line for lunch, during transition times in class, etc. Observation notes included statements that “the vast majority of verbalization occurs during breaks and snack” and that “most verbalization [for ELL] students occurred in non-instructional times.” Such comments indicate that ELL students’ opportunities to engage in interactions that push their output of
academic language appear to be limited in the mainstream classrooms. Consequently, the development of cognitive academic language proficiency (CALP) for these beginning level students may be hindered rather than helped by their participation in the mainstream.

**Attempted Initiation.** I chose to label this category *Attempted Initiation* rather than *Initiation* because I wanted to make clear that attempts by the ELLs in this study to initiate interaction were not always successful. Observers noted several examples of when the ELL students attempted to interact with their native English-speaking peers, only to have their attempts met with lack of patience or dismissal by their English-speaking classmates. Again, I cite an instance with Bluh.

While writing, subject asked a native English speaker how to spell the word ‘last’. After the native speaker asked to repeat the question multiple times the NS finally said “I just don’t understand you.” Subject gave up and raised his hand for teacher assistance.

While some students attempted to be helpful in the various classrooms, it was frequently observed that the native English speakers in the classrooms lacked patience with the beginning level ELLs and sometimes treated them with little respect.

Ben, a sixth grade level 1 refugee from Malawi, was attempting to work on a science activity with students at his table group when a boy laughs at him when he writes the wrong letter while a girl spells a word for him to write. Both members of his group (boy and girl) seem to treat him condescendingly. Ben puts his head down on the desk.

Ben was clearly upset and angry, as this was not the first time that he was observed to have difficulties in the class, especially with this particular group. Noticing the difficulties that Ben
was having with students at this table, the teacher changed the seating arrangements the following day. In the follow up debriefing session with Ben, he expressed that he often felt “angry” when he was in his mainstream classroom, confirming what I had observed throughout the observations. His interactions and attempted initiation of interactions were observed to be much more positive in the previous observations in the ELL sheltered classroom.

It was interesting to note that former ELL students in the classrooms tended to be much more patient with their beginning level ELL classmates. Their willingness to help and to respond to the new ELLs’ initiated interactions was quite apparent. Although it is often common for former ELLs to try to distance themselves from the ELL label and to try to fit in, the former ELLs in this study seemed much more likely to extend assistance and exhibit patience to their beginning ELL classmates even when they did not speak the same native language. The bond they shared living in a new country and learning a new language may have been an important factor. As a result, former ELLs were willing to help and beginning ELLs were comfortable initiating interactions.

Behavior Problems. The mainstream notes actually had no occurrence of comments identifying the beginning ELLs in the study as having been behavior problems in class. In fact, during follow-up debriefings with each teacher, it was frequently mentioned how much they enjoyed having ELL students in their classrooms because they rarely caused problems. As one teacher stated, “I figured they'd be bored and act out, but I've never seen a behavior issue with an ELL student.” Another teacher shared that sometimes the ELLs have behavior issues “for no other reason than they don't feel like they can participate, so they get squirrly.”

What the mainstream observation notes did reveal were numerous mentions of behavior problems with the non-ELL students in the classrooms. This was also a frequent topic brought
up by the mainstream teachers during the debriefing sessions and during informal exchanges when students were out of the room. After one observation, the well-seasoned teacher felt the need to justify to me the arrangement of the desks into rows, expressing that, while not his preferred arrangement, this was done as a means of controlling the “very difficult behavior issues in this class.” On another occasion, he also commented that student behaviors had become increasingly worse over the past few years, this year being the worst, and that he felt that the highly structured environment that he maintained in the classroom provided a “safe, predictable, and calm atmosphere, often absent in these students’ home lives.” His strategy did appear to work in keeping students focused and in minimizing disruptions. I was impressed by how eager the class was, as a whole, to answer questions, and in particular to use the white boards during math review sessions. Unfortunately, the rows of desks and highly structured, primarily teacher fronted classroom did not leave much room for the ELLs to engage in large amounts of peer interaction.

**Positives.** I have included *Positives* as its own category because, quite frankly, there were so very few positives noted in the observers’ notes. These positive comments tended to be primarily regarding the set up of the classrooms (labeling, visuals, table groupings) or evidence of the use of GLAD strategies in one of the classes. There was one very notable exception.

*The Exceptional Day.* On Friday, the last day of Benjamin’s mainstream observations, the classroom teacher made several significant changes to try to engage Ben. The teacher had commented that Ben had been struggling a lot, not just academically but socially as well during the current school year. (The teacher had moved up with the same group of students from the previous year. So this was his second year with Ben.) The teacher also noted that Ben was having an especially difficult week.
The previous day, Thursday, proved to be the worst for Ben. He had been shut out and ridiculed by the other students. In one instance, during a science experiment, his group would not allow him to participate. He was visibly angry and eventually became completely disengaged, putting his head down on the desk.

Friday was much better. That morning the teacher had the students change their seat assignments, which put Ben with a new group of students. The teacher’s decision to make the changes on that particular day had everything to do with trying to improve the situation for Ben, as was what followed.

The teacher had another Friday surprise for the class. He had been contacted by a Peace Corps volunteer in Africa who asked if he would be interested in having his class serve as pen pals with her group of 11 and 12 year old students in a refugee camp in Malawi. Benjamin’s teacher later shared with me that in the past he had always turned down these sorts of requests, finding it difficult to make time for such activities since it was not part of the regular curriculum and took time away from instruction. He was about to turn this one down as well until he started to think about the situation with Benjamin. Because this request was from a teacher in a refugee camp in the same country where Benjamin had also been a refugee just over a year ago, he decided that this might be a good opportunity to better include Benjamin in the classroom community.

Thus, at 10:20 on this Friday morning, just as the first snow of the year was beginning to fall outside the window (an event that could not go unnoticed by a class of 6th graders eager for today’s early release), the teacher announced the project, telling the students that they would each be exchanging letters with students in Malawi. He reminded the class that this was Benjamin’s home country and pointed it out on the map. Then, as he glanced toward the heavily
falling snow just outside the window, he asked Ben if the kids in Malawi have snow too. Here’s how I documented Ben’s response:

Ben smiles and responds in full voice, “They don’t know snow.” Ben is becoming the authority. He is very excited about this project. [As the students begin writing their letters introducing themselves to their future pen pal], the teacher has Ben come to his desk and asks him what he wants to say in his letter. Teacher writes while Ben dictates.

During this writing activity, students in the class seemed to see Ben in a very different light, asking him lots of questions about Malawi and about the weather and the animals there. They also asked him about coming to the United States and how long it took to travel. It was interesting to note that I overheard the other ELLs in the class also sharing about their own journeys and the countries they came from. It was as if someone had suddenly given them license to be different. Some of the non-ELL students expressed surprise that some of their classmates spoke another language and had lived in another country. As I noted in the observation data, “Much more talking for all ELLs and exited ELLs in class today.” It was a rare positive day.

**Sheltered Findings**

The analysis of the sheltered observation notes revealed student behaviors far different from what was noted for these same students in the mainstream. What typified the sheltered observations were descriptions of jubilant, talkative and engaged students who were interacting and learning using their new language. As one observer noted, “Everyone is participating. Kids are motivated.” This is a sharp contrast to the largely negative experiences noted by observers in the mainstream classrooms.
Following the same process of analysis as previously outlined, the observation notes of the sheltered ELL classrooms revealed the following seven categories: Engaged, Excited/Eager, Initiated, Playing/Off-Task, Disengaged, Other Positives, and Other Negatives. Whenever they arose, I named the categories such that they would align with similar or related categories found in the mainstream notes. I found that two dominant categories emerged from the sheltered observations, reflecting more than half of the coded text in the sheltered notes. The categories of Engaged and Initiated Interactions each accounted for 26% of the observers’ comments.

**Attentive/Engaged.** In sharp contrast to the rival category of Disengaged/Alone, which was so dominant in the mainstream notes, the seven beginning level ELLs were observed here to be very much a part of the linguistic and content learning taking place in the sheltered classrooms. As I read through and coded the observers’ notes describing this high level of engagement, I found that the observers had used a wide variety of descriptors to capture what they were seeing in the classrooms. Grounded in these many descriptors, the codes that I applied to the sections of texts, which would eventually make up this category, included “engaged”, “attentive”, “intent”, “focused”, and “interested”.

Along with high levels of engagement comes increased willingness by the students to express their opinions, take leadership in the group, and even oppose the views of other students. Benjamin, who would later be observed to be frequently disengaged and angry in the mainstream classroom during the fall, appeared quite comfortable engaging in debate in this sheltered setting. According to the observer in his class, “In Ben’s group, they are arguing about if the peanut is from a tree or not, but he does not think so.” The observer noting this had clearly witnessed Ben’s willingness to engage in the debate and express his opinion even though it was not
supported by the group. The observer later notes, “Ben is so excited about his right answer even though the girls did not agree with him.”

Another student, Nadia, exhibited similar leadership in connection with her high levels of engagement. The observer noted, “Even though she was not in class yesterday, she shows interest … She is focused in her work group and is the leader who talks for the class.” Although Nadia was identified by both her mainstream teacher and the ELL specialist in her school as being the exception rather than the norm for a beginning ELL (because she was talkative and outgoing even though she was new to the country), her mainstream observations showed little of the confidence and leadership that was evident in the sheltered notes and video. What I observed from Nadia in the mainstream was a great deal of social discourse. Still highly social and often off-task in her sheltered classroom as well, Nadia appeared to be more willing to take risks with the development of academic language and leadership.

For a student in the United States less than three months, Hso was observed to be highly engaged in the sheltered classroom. I noted during one of this sheltered observations that “Hso is not verbalizing much yet, but is clearly on task and engaged with nonverbal interaction, much more so than when I observed him in the mainstream. He is very intent on labeling his map.”

**Initiated Interactions.** Sharing the top spotlight with *Attentive/Engaged*, as the most frequently occurring, is the category of *Initiated Interactions*. Unlike in the mainstream, where students’ attempts to initiate interaction were often dismissed or even ridiculed by classmates, in the sheltered classroom, students’ attempts to initiate were met with success. This is likely the reason that such initiations were so frequently noted by the observers. Successful initiation leads to more successful initiation.
To initiate interaction requires a foundation of confidence, engagement, and an expectation that the initiated interaction will be reciprocated. This necessary combination is clearly in place in the sheltered classroom, where students are interacting with other English language learners. However, when students are in the mainstream, one or more of these are absent, thus destabilizing the foundation and causing beginning level ELLs to be hesitant to take the step to initiate interaction.

**Excited/Eager.** Closely related to engagement is the category of *Excited/Eager.* Rarely identified in the mainstream, observers mentioned students being “excited” or “eager” frequently in their notes from the sheltered classroom. I narrowed down the numerous descriptors used by the observers to seven working codes: “excited”, “eager”, “energetic”, “active”, “confident”, “enthusiastic”, and “motivated”.

Several students stand out in this category as having had very positive experiences. I highlight Benjamin because he was arguably the most eager of the students, according to observer notes. Each day of the observation, the observer noted something, often several things, that indicated that Ben was very excited and motivated to be in class. On one such day, the observer noted, “he is in a hurry to ask what he can do and work with group members.” On another occasion, it was noted that he was “eager to participate during TPR.”

Ben’s excitement over the summer program and his participation in a sheltered classroom was also evident as we walked down the hall to participate in the debriefing session five months later. As we walked, he asked me what we were going to be doing. When I told him I wanted to talk to him about last summer’s language camp, he became very excited, smiling from ear to ear. He went on to ask if he could go to language camp again this coming summer. He was very happy to find out that the answer was “yes.” During the debriefing session, with the presence of
other students and the interpreter, Ben also asked if we could have a special class just for English after school. It was clear that his sheltered experience was indeed “exciting.”

Poh Reh and Sitara were also observed showing excitement during their experience in the sheltered classroom. Sitara was described as being “very active and not afraid to talk.” It was clear in the mainstream observations that both of these students very much wanted to be involved in the learning taking place. Both of their mainstream teachers recognized them as very bright. However, it was also very clear that the mainstream environment did not offer instruction at a level that was appropriate to their prior educational backgrounds and, most especially, their levels of language acquisition. In the sheltered classroom, the learning appeared to be much more accessible and appropriate to their needs, thus they could be openly “excited” about participating.

Even Christelle and Bluh, who were highly disengaged in the mainstream and somewhat more melancholy by nature, proved to be far more eager to participate in the sheltered classrooms. Although Bluh was sometimes identified as being unsure about the task, the observer noted, “when the directions were restated, Bluh eagerly began his work and went to the window sill to observe his experiment.” This observation is in stark contrast to descriptions of Bluh sitting ‘stoic and frozen at this desk’. Bluh is now part of the whirlwind rather than being surrounded by it.

**Playing/Off-Task.** It is not unusual for students, especially ten to twelve year old students to occasionally be off-task or playing during class. It was, therefore, not surprising to find such statements in the sheltered classroom notes. However, in contrast to the notes from the mainstream classroom, the statements about off-task behavior were stand-alone statements. In the mainstream observations, the students were seen as being isolated and alone while they
played with things in their desk or when Ben was “playing with a snack wrapper.” When students in the sheltered classrooms played or were off-task, the notes indicated and the video confirmed, that the students were typically with another student, generally with smiles on their faces. They tended to play for playing’s sake, not in connection with being alone or disengaged.

**Disengaged.** Despite many positive elements observed in the sheltered classroom, students also showed evidence of disengagement though with much less frequency than the mainstream classroom. One student in particular, whom I previously described as somewhat melancholy by nature, accounted for nearly all of the references for disengagement. In one observation, she was described in the notes as “passive today”, “distracted” or “seem(ing) to be bored.” On another day, “Christelle is isolated while her group is working at the board. … Christelle has an indifferent attitude today.” Nevertheless, Christelle was much more eager and engaged in the sheltered classroom than she was during the observations in the mainstream classroom, with the observer noting, “Christelle is very proactive.” As I reviewed video of Christelle after one observations, I had noted, “I am surprised to see how often Christelle raises her hand to volunteer answers. So different from in the mainstream.”

Video clips and observation notes reveal that disengagement and playing/off-task were more prevalent when the pre-service teachers, as opposed to the certified ELL teachers, were in charge of the lessons. This was particularly true when the lessons were not well planned and did not account for varying levels of language ability in the classrooms. This was the case both when tasks were too difficult and when they were too easy, which resulted in observers’ noted student boredom. This appeared to be the case with Christelle and Nadia, who were in the same classroom (observations of each occurred on different days).
Extending a Helping Hand

As I reviewed the quotes from the sheltered notes, I found an additional, previously overlooked category. It became quite apparent that the ELL students showed a willingness to help each other in the sheltered setting that was not shown in the mainstream classroom. This willingness to help is evident in observer notes such as “Ben was concerned about others who weren’t participating” and “Christelle helped her peers in the game” or when Ben was observed to be “willing to explain to his group member… and [did] actions to help her understand.”

When students are engaged, confident, eager, and comfortable, they appear to be willing to express concern and extend a helping hand to their classmates. In doing so, they are not only showing compassion, but they are increasing their own engagement and creating their own opportunities to interact.

Qualitatively Opposed: A Summary

To bring the highly contrasting qualitative findings into clearer focus, I decided to employ one more analytical device. Figure 14 summarizes the number of coded comments I considered positive and how many I considered negative for each of the two settings. The criteria I used to determine whether a particular comment fit into either category was based on two key questions: 1) Did the comment exhibit a raising or lowering of the affective filter (anxiety) for the student? or 2) If the comment pertained to classroom environment, did the comment illustrate what is considered to be a best practice for working with English language learners?

Supporting what has already been described in the preceding sections, Figure 14 illustrates a nearly complete inverse between the sheltered and mainstream classrooms. Whereas
the sheltered notes indicated 71% positive comments and 29% negative, the mainstream notes indicated only 15% positive and 85% negative.

![Figure 14. Positive and Negative Categorization of Coded Observation Notes](image)

In the preceding section on mainstream findings, I used a Wordle illustration to show the dominance of the word *assessment* in the observers’ notes. At this point, I would like to sum up the findings for the qualitative component of this study by sharing two Wordles, one from the mainstream notes (Figure 15) and the other from the sheltered notes (Figure 16). Although I’ve used the same coded text to generate the new mainstream visual, I have deleted the word *assessment* from the mainstream display. This deletion served to 1) allow the other dominant words and phrases in the mainstream comments to become more salient; and 2) allow for a more equitable comparison between the two settings, because formal assessments were not required nor present in the sheltered classrooms of this study. Figures 15 and 16 capture the essence of the qualitative component of this mixed methods study.
Figure 15. Wordle of Mainstream Coded Notes Excluding "Assessment"

Figure 16. Wordle of Sheltered Coded Notes
Chapter 6

Discussion and Conclusions

Given that the importance of interaction has been firmly established in SLA research, the focus of many researchers over the past three decades has shifted to further analyzing the components of interaction in more detail. In this study, however, I have stepped back to ask a more fundamental question of how often beginning level English language learners actually interact using English in the mainstream classrooms compared to the sheltered instructional settings of the sheltered ELL classroom and the ELL pullout group. I then examined how qualitative factors with the mainstream and sheltered classrooms served to heighten or suppress interaction in these settings. The overarching question tying both the qualitative and quantitative components is which instructional setting provides the optimal environment for beginning level English language learners to interact and engage in using the target language?

The results of this study, involving English language learners in the upper elementary grades, provides a combination of quantitative and qualitative support for previous qualitative research on interaction between mainstream and sheltered settings. These previous studies, almost exclusively with middle school and high school students, found that English language learners, even at advanced levels, interacted far less in the mainstream than in their ELL classrooms and were additionally found to be more anxious and less engaged than in their sheltered ELL classes (Duff, 2001; Harklau, 1994; Miller, 2000; Pappamihiel, 2002). The beginning level elementary learners in this study were likewise shown to be largely disengaged in the mainstream classrooms.

In the quantitative component of the study, I compared the frequency of English verbal interactions for 3rd - 6th grade beginning level English language learners across three different
instructional settings. The purpose of the quantitative component was to document if substantial differences in interaction rates existed between the mainstream classroom and ELL specific instructional settings, namely the ELL sheltered classroom and the ELL pullout group. Further quantitative analysis then sought to establish whether these differences were attributable primarily to the ELL students’ interactions with teachers or with classroom peers within each setting.

I further investigated, through qualitative analysis of video and observation notes, what environmental factors within the mainstream and sheltered classrooms may have contributed to differences in interaction frequency. Such understanding of interaction frequency and the factors that contribute to English interaction in different instructional settings affords insight into which setting offers the optimal environment for verbal interaction to take place in the target language.

Vygotsky’s social development theory (Vygotsky & Luria, 1934/1994), which emphasized the importance of social interaction in linguistic and cognitive development, serves as the cornerstone for the premise of this study. More recent research, conducted over the past thirty years, has established interaction as an essential component of second language acquisition. This is amply supported particularly with respect to pushed output and negotiation (e.g. Long, 1981; Pica, 1994; Swain, 1985; Varonis & Gass, 1985). It has been shown over and again that the ancient Chinese adage, “Tell me and I'll forget; Show me and I may remember; Involve me and I'll understand,” is as true in language learning as it is in any other educational endeavor.

The quantitative results of this study show that beginning level ELLs exhibited a significantly higher frequency of English verbal interactions (an average of seven times higher) when they were with other ELL students in a small pullout group and nearly three times more
often when they were in an ELL sheltered classroom with their English language learning peers as compared to when they were in regular mainstream classrooms. Such findings are similar to those discussed in a qualitative study of high school English language learners in Australia who, upon transitioning into a mainstream high school, “stopped trying to speak English” (Miller, 2000, p. 96). Studies by Duff (2001) and Harklau (1994) offered similar findings of reduced interaction in mainstream classrooms.

Since the question of whether interaction is a necessary component of language learning has already been firmly established, it follows that reduced levels of interaction frequency, like those found in the mainstream classrooms of this study, would then negatively affect the rate that students are able to acquire language. And, if Vygotsky (1997) is correct that cognitive development and learning result “when the child is interacting with people in his environment and in cooperation with his peers” (p. 35), then students’ lack of interaction in any given environment could likewise negatively affect cognitive development and learning in a broader sense.

As I further investigated, through qualitative analysis, the behavioral and environmental factors that may have influenced the significant differences in interaction rates as they were documented by observers, I found that the ELL students were highly disengaged in the mainstream, whereas they displayed high levels of engagement and interest in the sheltered ELL classrooms. These findings concur with other studies involving older language learners at higher levels of acquisition. Harklau’s (1994) ethnographic research with high school ELLs, similarly found that “students were often withdrawn and noninteractive in mainstream classes (p. 252). Pappamihiel (2002), who investigated language anxiety levels for 178 middle school ELLs in
mainstream and sheltered classrooms, attributes disengagement and withdrawal from learning to high levels of anxiety in the mainstream classroom.

Anxiety for the students in this study was indeed quite evident in the observations for the mainstream. The student’s lack of exposure to English prior to being placed in the mainstream 3rd-6th grade classrooms meant that they were ill equipped to engage in even the most basic grade level activities in the classroom. Six of the seven mainstream teachers in this study shared that they had had virtually no training for working with the English language learners in their classrooms, let alone those with limited or no formal schooling. This lack of ELL training for the teachers is even lower than the four hours of training reported by Crawford and Krashen (2007) for the average teacher working with ELLs in U.S. schools. This was true even for the student teacher in this study just finishing his teacher training program, who shared that the only mention of ELLs in his teacher preparation courses occurred briefly during one lecture in a special education course. Therefore, even if there were no other factors that contributed to increasing anxiety for these students, it is clear that the teachers were not prepared to deal with language anxiety with their beginning level ELL students as they were ill equipped to adequately meet their learning needs and effectively engage these students.

Other factors were, in fact, present. One of the factors that visibly increased anxiety and caused students to be disengaged to a much greater degree in the mainstream was the extensive emphasis on formal assessments. Since new ELL students are required to be given a qualifying/placement test within ten days of first attending a U.S. school, they are, from the very beginning, faced with having to take tests for which they have no possibility of being successful from the very first. In addition, all ELLs, even if they have only been attending a U.S. school for a few days and have never attended school in their home country, are required to take the math
and science sections of the statewide annual assessments. Students who are faced with repeated failures and an inability to even attempt these assessments may not only develop anxiety about assessment but also develop negative attitudes and anxiety towards school in general.

As most of the students in this study had very little experience, and in two cases, no experience, attending school prior to coming to their current school, they lacked, not only the language skills necessary to be engaged in and be successful in the mainstream, but also the background in content and the literacy skills in their first language that would have helped transfer knowledge into their new language of English. The stress of competing and being assessed on knowledge that they have not yet been taught or for which they do not have the skills to access raises what Krashen termed the “affective filter” and further increases their anxiety in the classroom, leading to disengagement and withdrawal from the learning process. The instruction and challenges of the mainstream classroom for these beginning level students is rarely within the zone of proximal development discussed by Vygotsky (1997) and is far above the i + 1 that Krashen (1982) recommends for acquisition to occur.

The sheltered settings, offered a distinct contrast both in quantity of interaction and in observed behaviors and engagement in learning. So, what was it about the sheltered settings, both in the large classes of the summer program and in the pullout groups that contributed to the marked differences that were observed?

Three factors highly present in the sheltered settings contributed to these significant differences: 1) the marriage of content with a distinct focus on language development, 2) instruction appropriate to each student’s background and level of acquisition, and 3) teachers with significant ELL training and background. Although Guided Language Acquisition Design (GLAD), a popular in-service training in this district as well as in this state, does a good job of
marrying language and content, in the one mainstream classroom where it was observed to be employed, the beginning level ELLs continued to be disengaged. In the sheltered classrooms of the summer program, in which GLAD strategies were frequently employed, the same students were highly engaged. Although it appears that the combination of language and content is important, it clearly takes more to encourage interaction for beginning level students. The instruction needs to be at the appropriate level, based on students’ background and level of acquisition for the students to be engaged. For this to occur, teachers need to be able to first understand what the language abilities and needs of the students are and then be able to design instruction that meets those needs. For a multi-leveled classroom, as were all of the classrooms of the summer program, this can of course be a challenge. However, when teachers are well trained and experienced in working with beginning level pre-literate students in multi-level settings, they are more likely to deliver content lessons that plan for language learning at levels appropriate to the needs of beginning level students. This contributes to the increased engagement found in these settings.

Although the preceding three factors clearly played a role in increasing interaction and lowering anxiety, the most influential factor in both of the sheltered settings appeared to be the fact that students were interacting with other English language learners who shared a common experience, even when they were not from the same language background or at the same level of acquisition. Mackey (2012) identified interaction as “dynamically shaped by participants’ expectations, experiences, and beliefs about their interlocutors and the communicative situations in which they find themselves” (p. 24). It was observed to be the case in this study as well that student experiences and peer background are important not only in shapings interaction, as Mackey contends, but additionally, in determining to what extent interactions will occur at all.
When students were surrounded by native English speakers and additionally required to engage in academic rather than social language, students tended to withdraw. Pappamihiel (2002) described interactions of middle school ELLs with native English speaking classmates as “strained and avoided” (p. 340), as students themselves indicated that their levels of “interactional anxiety” were higher than their “academic anxiety” in the mainstream (p. 344). Because beginning level students were engaged entirely with other English language learners when they were in sheltered classrooms, they were more comfortable engaging in interactions. They did not perceive themselves as having to compete with native speakers academically or interactionally.

One question that may arise in regard to the students’ use of academic or non-academic language is whether or not the mainstream teachers afforded sufficient opportunities for small group or paired activities. These are often encouraged in pre-service and in-service trainings to help make content comprehensible and increase academic verbal interaction. Even if teachers provide these opportunities, another question arises as to whether the ELL students actually engage in the small group or paired opportunities.

Although a separate study would provide more definitive evidence for how students interact in small groups with native speakers, when mainstream teachers in this study were observed to employ such strategies as having students work in small groups or in pairs with native English speakers, some students actually exhibited less interaction than when they were not in small groups. ELL students were often observed to have been either pushed aside by the English-only students and not included in the small group interactions or the ELLs neglected to interact, allowing their native English speaking peers to take over the activity.
The observational scenario describe as the “exceptional case” in the chapter 4 exemplifies such a case where an ELL showed much reduced verbal interaction when placed in small group or when paired with a native speaker. In the case of Poh Reh, his normally high rate of interaction in the classroom was completely shut down in the paired activity. Varonis and Gass (1985) suggest that interaction between native speakers and non-native speakers “actually discourages negotiation” because of the “inequality in the (language) status of the participants” (p. 86). They contend that the higher rates of negotiation between language learning peers (NNS-NNS) in their study to the fact that the individuals in the pairs were both non-native speakers and were therefore more comfortable admitting and recognizing miscommunications because of their “shared incompetence” (p. 84). They further argued that non-native speakers “do not lose face by negotiating meaning [with other non-native speakers] in the same way they might with native speakers” (p. 85). For Poh Reh, and the other beginning level students in this study, to be open to taking risks by engaging in the use of academic language and exposing themselves to ridicule, whether real or perceived, may be too much to expect.

**Conclusions**

I conclude that the mainstream grade level classroom is not the optimal instructional setting for beginning level English language learners to engage in meaningful verbal interactions which are requisite for language development and the literacy skills needed for future academic success. It is clear that for most ELL students being surrounded by native English speakers in mainstream classrooms actually serves to inhibit student production in the second language rather than promote it. The exclusion from meaningful interaction for all the reasons discussed thus far contributes to increased anxiety in the classroom and affects both engagement and learning. The sheltered ELL classroom, composed of English language learning peers, offers an
optimal environment for both increased verbal interaction and low levels of anxiety. Although pullout groups may not be the most desirable method for delivering instruction, due to the infrequency of student contact during the school week, they do provide the best opportunity for beginning level students to interact when the only option is a mainstream classroom. As stated in the 2011 Washington Transitional Bilingual Instructional Program’s Report to the Legislature,

> When students with little or no previous exposure to the English language enter the public school system, they are most often unable to fully benefit from instruction in English and may experience a high risk of academic failure (Malagon, McCold, & Hernandez, 2011b, p. 13).

Beginning level students first entering school in the upper elementary grades need instruction that develops both their language skills and builds content and literacy skills in English at a level that is appropriate to their existing level of understanding. Without this foundation in place, students will continue to struggle with academic success even beyond the first few years.

**Limitations**

Certainly every classroom and every study will have a different situation and a unique set of challenges. This particular study was no exception. While efforts were made to schedule observations around anything that would be perceived as outside the norm of daily classroom activities, it must be acknowledged that scheduling “normal” mainstream observations that did not include some sort of district mandated assessment or practice for annual state assessments proved to be virtually impossible. Mandatory assessments and direct test preparation have clearly become part of the normal daily procedures of the mainstream classroom. As these disruptions in instructional flow constitute the reality of the mainstream classroom, I hesitate to
classify them as limitations of the study. The constant testing and test preparation that have been forced by punitive accountability measures in No Child Left Behind are currently, much to the beginning ELLs detriment, part of the fabric of the mainstream classroom environment. As one of the mainstream classroom teachers so colorfully commented in regard to the testing of all students: “We are constantly weighing the pig and never feeding it.” This sentiment has been found to be shared by many teachers, whether they teach in the mainstream classroom or serve as ELL specialists.

Because the ELL sheltered classrooms in this study were part of an optional summer program, the emphasis on making adequate yearly progress (AYP) and achieving predetermined benchmarks were not the driving force that they were in the mainstream classrooms during the regular school year. The language camp teachers, and the program as a whole, were able to focus on taking students from wherever they were in the language development continuum and move them forward. The overall atmosphere of the summer program was therefore more relaxed in this regard.

Another limitation for this study is that there existed a relatively small sample size of only seven students who had participated in all three settings. This did not affect the statistical analysis for the quantitative component of the study, however, given that the r values were quite high, between .80 and .91 for all aspects of the analysis. Additional quantitative studies of this sort with a larger sample size would be welcomed to further verify the reliability of the quantitative findings.

**Future Studies**

Whereas previous studies have been conducted during the regular school year at the middle school and high school levels (Duff, 2001; Harklau, 1994; Miller, 2000; Pappamihiel,
2002), it is rare to find the same elementary students participating in both mainstream and sheltered settings during the regular school year. The difficulty in conducting such studies is that few content-based, truly sheltered programs exist in the elementary grades. Additional studies that compare student interactions in elementary sheltered and mainstream classrooms during the regular school year would be most useful.

Since this study did not specifically set out to document academic v. non-academic interactions, an additional study along these lines would be most helpful in establishing to what extent such differences may exist.

A Final Word

As more and more schools move toward considering immediate and full mainstreaming of their beginning level ELL students into grade level classrooms, administrators and law makers alike need to consider that by not allowing students access to instruction appropriate to their academic and linguistic needs and by not allowing them participation in classrooms that provide low anxiety and adequate opportunities for meaningful peer interaction, they may actually be slowing ELL students’ progress and delaying or even preventing them from achieving at or above grade level on mandatory assessments as they move through the system. With the new highly challenging Common Core State Standards in place and the Smarter Balance assessments soon to be implemented, we need to ensure that the ELL students in our schools have a solid foundation of language upon which to stand if they are to be successful in the future.

We must take care that we are not turning back the clock to pre 1974. It was then that the case of Lau v. Nichols was heard and Supreme Court Justice Douglas delivered the opinion of the Court, stating:
We know that those who do not understand English are certain to find their classroom experiences wholly incomprehensible and in no way meaningful … Where inability to speak and understand the English language excludes national origin-minority group children from effective participation in the educational program offered by a school district, the district must take affirmative steps to rectify the language deficiency in order to open its instructional program to these students….There is no equality of treatment by providing students with the same facilities, textbooks, teachers and curriculum, for students who do not understand English are effectively foreclosed from any meaningful education (n.p.).

We must ensure that we are providing the best possible opportunity for all students to succeed. And, for English language learners, we must ensure that we are not making the mistake of submersing rather than immersing them in language.
Bibliography


Lau v. Nichols, 72-6520 (Supreme Court of the United States January 21, 1974).


Figure 17. Wordle of Strategies Found in CALLA, SIOP, and GLAD
Appendix B

Research Observer Equipment/Materials Protocol

1. Each observer will be provided with one digital Flip camera, a tabletop tripod, and observation recording sheets. (A back up camera is available from Joan if needed.)

2. At the end of each observation day, return the packet containing the camera equipment along with the completed observation sheets for that day.

3. Since the IRB approval requires that I keep all data secured, I will download the video to my computer and make back up disks, which will be stored in a secure location. Please note that observers will not be downloading any files to their personal computers.

4. Before the next day’s observations, pick up the observer’s packet. Batteries will all be recharged and memory cards will be empty and ready for more data.

5. Remember that confidentiality needs to be maintained. The IRB approval for this research requires that students’ and teachers’ names not be shared with anyone outside of the research team.
Table 9

Overview of Data Collection

<table>
<thead>
<tr>
<th>Setting</th>
<th>Mean Class Size</th>
<th># of Students Observed</th>
<th># of Observations</th>
<th>Average # of Observations per Student</th>
<th>Total Time Observed</th>
<th>Average Observation Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainstream</td>
<td>22.5</td>
<td>13</td>
<td>40</td>
<td>3.1</td>
<td>4657 min.</td>
<td>116 min.</td>
</tr>
<tr>
<td>Pullout</td>
<td>2.5</td>
<td>17</td>
<td>38</td>
<td>2.2</td>
<td>1386 min.</td>
<td>36 min.</td>
</tr>
<tr>
<td>Sheltered</td>
<td>34.4</td>
<td>8</td>
<td>23</td>
<td>2.9</td>
<td>2042 min.</td>
<td>89 min.</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>20*</td>
<td>101</td>
<td>8.2</td>
<td>8085 min.</td>
<td>80.3 min.</td>
</tr>
<tr>
<td>3 Settings**</td>
<td></td>
<td>7</td>
<td>63</td>
<td>9</td>
<td>5086 min.</td>
<td>80.7 min.</td>
</tr>
</tbody>
</table>

*Totals represent separate individuals. **Same students in all three settings.