The Dissertation Committee for Lisa Marie Volle Certifies that this is the approved version of the following dissertation:

Virtually Speaking: Comparing Foreign Language Speaking Performances in Distance Education and Face to Face Classes

Committee:

Zsuzsanna I. Abrams, Supervisor

Elaine K. Horwitz

Zena T. Moore

Paul E. Resta

Diane L. Schallert

Virtually Speaking: Comparing Foreign Language Speaking Performances in Distance Education and Face to Face Classes

by

Lisa Marie Volle, B.A., M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August, 2010

Dedication

For Sara

~

In Memory of students from this study: Carolyn and Scott

Acknowledgements

First, I would like to thank the faculty members who served on my committee. Dr. Zsuzsanna Abrams came to UT like a breath of fresh air. I was very privileged to take two classes with her. She mentored and encouraged me to write my first publication. Dr. Zena Moore contributed a great deal to the FLE department and I took her many classes – from teaching culture to assessments. We spent a lot of time together before her UT retirement, and I appreciate her very much. If I describe Drs. Abrams and Moore as the sails, Dr. Elaine Horwitz is the anchor. Her depth and breadth of knowledge in second language acquisition and ability to motivate one to look closer at the discipline's issues is truly appreciated. Dr. Paul Resta was a great model for online teaching with high standards, quality and content. His course, on collaborative online learning, was an eye opener and provided me with the student's view of online learning. Dr. Dianne Schallert's energy and enthusiasm were unparalleled. Each provided me with views into the vast field of education and into foreign language education. The knowledge and experiences I gained by being involved with the faculty members on my committee will forever aide me in my journey.

I would also like to thank Kathy Sowash and Alicia Thomas in providing so much support as I worked through my program. Kathy, especially, has become a dear friend. I also want to thank Barbara Haley and Katherine Pangelinan at Central Texas College

who truly made my life easier throughout this process. Their support in my teaching departments and as personal friends has meant a great deal to me. There are many from CTC I would like to thank for their encouragement and support: Dr. Maria Bravo, Dr. Joyce Bateman-Jones, Dr. John Henderson, Dr. Jerry Herrick, Dean Don Mikles, Sharon Collins, Herb Maglietta, Dr. Dana Watson, Dr. James Anderson, Dolly Kinder, Chris Mott, Barbara Merlo, Patricia Fasano, Kathy Cotter, Tracy Cook, Michelle Anderson, Jan Anderson, Anna Holston and Dr. Suzanne Morales-Vale.

I would also like to acknowledge the support of colleagues and friends from the Heinle, World Languages Consulting Group: Dr. LeeAnn Stone, Greg Rivera, Andrew Tabor, Nick Agnew, Dr. Mark Overstreet, and Edith Beard-Brady. Others from the publishing world include: Lisa Patterson, Janlyn Reed, Kristen Ciesielski, Amy Richards, Jay Hale, Ellen Edwards, Sherry Robertson, Richard Colangelo, Laura Wolfe, Sally Dennehy, Susan McElwain, LeDawn Webb, Maggie Miller and Janet Taborn.

The following friends provided much support and strength throughout my studies: Nouzha Swimlar, Carol Abbassi, Dr. Susan Pardue, Khedi Megateli, Dave Haley, Gorden and Brenda Tarnow, Mike and Juanita Bowles, Sandra and George Brown, Vanessa Sczcepanski, Elizabeth Guzman, Maria Volle and my friends on FaceBook and in Second Life. I especially want to remember, Dr. Leslie Jarmon who guided my learning in SL.

I owe a great deal to the students over the many years that have allowed me to get to know them and learn with and from them. I am especially appreciative of the students who participated in this study.

I also want to acknowledge my fellow UT graduate students with whom I have studied, presented papers, and have shared so much of this experience: Dr. Kari Lei, Dr. Fu-An Lin, Dr. Nihat Polat, Dr. Scott Gibby, Susan Valentine, Dr. Given Lee, Dr, Jin

Huh, and Dr. Michael Tallon. Michael has been a steadfast friend who has encouraged me at every step. I am thankful for his friendship.

Finally, I could not have succeeded in completing this dissertation without the unconditional support of my family: my parents, Adva J. Volle and Dr. Robert L. Volle, my siblings, L. Michelle Borden, Robert L. Volle, Jr., Dr. Steven J. Volle and John M. Volle, and my daughter, Sara Maria Lisette who is my center and *the wind beneath my feet*.

Virtually Speaking: Comparing Foreign Language Speaking Performances in Distance Education and Face to Face Classes

Publication No._____

Lisa Marie Volle, Ph.D.

The University of Texas at Austin, 2010

Supervisor: Zsuzsanna I. Abrams

The purpose of this study was to investigate the performances on speaking assessments in a first semester Spanish course to determine if there are differences among four sections based on required speaking tasks and mode of class delivery. Two of the classes were conducted online in a distance education delivery mode and two of the classes were conducted in the classroom. One section of delivery mode was required to fulfill speaking tasks throughout the semester while the alternate sections were offered the same speaking tasks as optional activities. Furthermore, the study investigated testing items repeated prompts versus new prompts to determine if there were differences in performance. Finally, the study investigated student performance confidence and beliefs about course materials.

The data were collected through quantitative methods (oral rubric scores) and qualitative methods (open-ended questions). A total of 64 students participated in the study. Eight students completed the distance education course with optional speaking

tasks. Thirteen students completed the distance education course with required speaking. Eighteen students completed the classroom course with optional speaking, and 24 students completed the classroom course with required speaking.

In general, all classes made improvement between the first assessment and the second assessment. However, the mean performance scores for three of the four speaking prompts did not show significant difference among the four classes. In addition, there were no significant differences in mean scores of performances at the introduction of the two new speaking tasks on the final assessment. Only with one speaking repeated measure was there a significant difference. The significant difference was between the classroom with required speaking section and the distance education with optional speaking section. With an analysis of the Target Language Use, it was determined that the differences were due to lexical complexity and morphological complexity.

In general, the self-report of confidence levels on the final oral assessment showed no significant differences among the four classes. There was a significant drop in confidence on one of the two new assessment prompts for all classes. Student comments indicated that there were frustrations with cognitive overload and dealing with ambiguity.

Table of Contents

List of Tables

List of Figures

List of Illustrations

Chapter 1: Introduction

The Background of the Study

With the rapid increase and growth of technologies, distance education courses have presented themselves as an alternative way to gain information and skills in lieu of the traditional classroom. In 2001, 90% of public 2-year institutions offered distance education courses (Tallent-Runnels, 2006). According to Pethokoukis (2003), online enrollment is increasing by 33% per year. In 2004, Miami-Dade Community College was ranked second in highest online enrollment with 57,026 students. The University of Phoenix, primarily an online institution, was ranked number one with 115,794 students (Aemmer, 2006). Teaching and learning with technology has become a growth industry and shows no signs of slowing down. Discussing today's distance education is done in terms of describing the fifth generation, even though we are moving with lightening speed into the future with "Digital Natives": learning on portable devices on the move, at anywhere and at any time.

The published research in distance education that concerns foreign/second language acquisition is lacking. In one recent theory-based meta-analysis of the empirical literature from 1985-2002 on undergraduate pedagogy in distance education, Lou, Bernard, and Abrami (2006) cite only two foreign/second language studies from a total of 103 studies.

Technological tools and courseware also supplement the traditional classroom or replace part(s) of the traditional classroom. Warshauer (1996), Kern (1995) and Swaffar et al. (1998) note that most foreign/second language research articles that concern

technology do so as an adjunct to the teaching classroom and not as the focus of the classroom itself.

Today, students enrolled in a hybrid/technology enhanced course may spend less time in the classroom with an instructor because certain parts of the curriculum can be performed "on-line" and/or at home. Even though computer-assisted language learning (CALL) has a much longer history of research than distance education, the two share much in their orientation. Abrioux (1989) was one of the first to notice the similarities. Because the learner and teacher are both separated in computer-assisted language learning activities and in distance education courses, providing students with immediate feedback is an advantage that CALL activities can bring to the distance education student. However, he only describes students benefiting primarily listening and writing-based CALL, or CALL incorporated activities. He writes, "Even if oral communication is generally more highly valued than written communication, the close relationship between the two facilitates mutual reinforcement" (Abrioux, 1989: 23). Since his writing 21 years ago, very few research articles address oral development in foreign language learning with technology.

In a review of literature from 1990-2000 in computer-based technology use in second-language learning, Liu, Moore, Graham, and Lee (2003) find that there were only six articles that investigated how technology could be used to promote speaking skills (Borrás, 1993; Coniam, 1998; Derwing et al., 2000, Gonzalez-Edfelt, 1990; Johnston & Milne, 1995; and Liaw, 1997, cited in Liu et al., 2003). Liaw (1997) described using computer books to help provide content and facilitate classroom conversation. Borrás (1993) and Johnston and Milne (1995) found that multimedia software increased

communicative exchanges. Coniam (1998) and Derwing et al. (2000) reported on the use of speech recognition software. The feedback provided was not reliable for nonnative speaker utterances. Liu et al. highlight "the need for more research in the less-explored skill areas such as speaking, listening, and culture" and a need for the research to "focus on … vocabulary acquisition and language production" (264).

In 2007, Chun collected the top ten most popular CALL topics and the number of articles that appeared in the two journals: *Language Learning & Technology* and *CALICO*. Between the years 2001-2006, she notes that Speaking ranks ninth. Between the two journals, 12 articles appeared. The speaking articles she cites are only about L2 pronunciation and intonation with software. She is optimistic about its possible future if software can provide meaningful feedback to learners about their waveforms and pitch contours (Chun, 2007).

In many foreign/modern language departments, "the traditional emphasis has been on literature rather than proficiency" (Schick & Nelson, 2001). Today's foreign language learner rarely takes a foreign language class to study its literature. It is much more likely to be because they want to "communicate" to speakers of that language, travel, understand another culture or be more marketable as a job candidate (McDonough, 2001). To further illustrate students' view of the value of speaking as an important goal, 76% of the students enrolled in a 2006 summer intensive Spanish course recently ranked speaking as the most important skill in a language learning class. The other 14% of the students ranked listening comprehension as the number one skill.

Many first year curricula are written, in response, with the goal of encouraging the student to express herself in the target language (TL). Met and Galloway (1996)

describe proficiency-based programs with communicative methodologies are those that "help students develop language skills that allow them to communicate their own thoughts and feelings in situations they are most likely to encounter. The emphasis is on what students can *do* with the language, as opposed to what they *know about the language*" (872).

Egan (1999) claims that "speaking is at the heart of second language learning but has been somewhat ignored in teaching and testing" (277). She provides several reasons for this. One reason is that speaking gives way to other curricular demands such as grammar. Another reason is that the teacher-student ratio makes the classroom logistics difficult. Speaking is also missing from student evaluations because it takes too much time to conduct and scoring the performances is not easily done objectively (Clifford, 1998 cited in Egan, 1999).

Today distance education students can work online as if they were sitting next to each other in the classroom. With the availability of audio and video, conferencing tools provide a type of face-to-face (f2f) context. However, there continues to be a paucity of foreign/second language research in one of the fastest growing and changing learning modes: distance education (Cahil & Catanzaro, 1997; Despain, 2003; Katoaka, 1987; Murphy, 2005; White, 1999). Furthermore, there is even less research literature that combines distance education with concerns of developing oral skills or the analysis of oral skills in a foreign language (Gamlin, 2005; Kötter, 1999; Rosell-Aguilar, 2005; Volle, 2005, Wang, 2004; Blake, 2008).

With the adult public demanding lifelong learning opportunities and the face of the online learner evolving, greater numbers of universities and colleges are under

pressure to offer a variety of courses in flexible formats. According to *T.H.E. Journal*, today there are more than 3,000 institutions of higher education involved in distance education/online courses. Distance Education courses benefit certain student populations in ways that traditional classrooms cannot accommodate. This stems from the ability of distance education courses to provide access to individuals who are "geographically isolated or restricted by work, schedules, and/or other considerations" (Ariza & Hancock, 2003). While it is clear that there is a number of benefits to the student, such as flexibility in structure, pace of learning, daily scheduling, immediate feedback on discrete point on-line activities, and exposure to a variety of authentic native speakers (Lively, 1997), it is still unclear whether the on-line learning context can have a similar impact on students' speaking performance to that of the classroom context.

Distance education language courses, if offered, should be equivalent to the traditional courses to offer similar speaking activities and assessments. The *ACTFL guidelines* (1997) urge language courses to include all language skills. Distance education courses in foreign languages should also be concerned with delivering instruction and allowing student practice that cover all aspects of foreign language learning. This includes reading, vocabulary, grammar, writing, listening, culture and the often under reported skill in course design, speaking (Volle, 2005). Therefore, this is the major impetus for the present study. Within communicative theoretical framework and in a context and practice based "classroom" with internet-based and off-line activities, Do distance education foreign language learners perform similarly to their in-class counterparts on the final speaking assessment? Is there a difference in speaking scores based on required speaking practice versus suggested speaking practice? Do students feel

that the course materials and technology available are effective in their developing oral skills?

THE PURPOSE OF THE STUDY

The overall purpose of the study was to investigate students' oral performance with different amounts of required speaking activities in two different modes of instruction: in-class instruction and online, distance education instruction. Although there is a wide range of meanings attached to the terms *in-class*, *online*, *technology enhanced* and *hybrid*, the following descriptions will serve to define this researcher's context. For the current study, the online students never met their instructor f2f (face-to-face), but interacted with her via the Internet. The in-class students had f2f access to the instructor during class and in office hours to include access by phone or email. Both online and in-class courses were technology enhanced. This required access to various CALL activities. For this study, a hybrid design – with partial in-class instruction/practice and partial online instruction/practice – was not employed. The research was carried out within the context of the individual instructor/researcher's institution – a community college with the main campus located in Central Texas.

Four sections of first semester elementary Spanish students, self-selected into one of two class sections or one of two online, distance education sections, were used in a convenience study. The present study explores the hypothesis that there are differences in performance on the final speaking assessment between the distance education sections and the in-class sections. A second hypothesis concerns the amount of required speaking. It is hypothesized that the sections with required speaking practice will perform better

than the sections that were given optional speaking practice. One section of distance education learners and one section of class learners were assigned required speaking tasks tied to their final grade throughout the semester. The alternate sections were given optional speaking practice (see Figure 1).

	Mode of Delivery	
attribute of oral task	Distance Education Section with required speaking practice	In-class Section with required speaking practice
	Distance Education Section without required speaking practice (optional)	In-class Section without required speaking practice (optional)

Figure 1: Four-way comparison of final speaking performance based on learning contexts

Language instructors are responsible for making more and more decisions in the design and course offerings and are also accountable for student progress and attrition rates. The findings of this study inform the SL practitioner with insight into the practice of student speaking in foreign language classes in distance education as well as in the class. It did not intend to identify what is wrong, but rather identify findings that suggest speaking tasks, techniques, and strategies to impact students' oral development. Finally this study informs curriculum design, instruction, and assessment.

The study is organized as follows. The current chapter serves as an introduction to the relevant issues. Chapter Two provides a review of literature of second language acquisition theory related to speaking, speaking in the foreign language curriculum, speaking with technology (on and offline), and assessing oral production. Chapter Two concludes with a description of the objectives and research questions of the study. Chapter Three presents a description of the institution, the student population and the

methodology. Chapter Four presents the findings of the student background, their current environmental context and the results of the data determined by the research questions. Chapter Five concludes with a discussion of the findings, the present study's limitations and suggestions for future research.

Chapter 2: The Review of Literature

INTRODUCTION

Describing the elements involved in how people "learn" a foreign or second language run along a continuum that involve many strings of research from the macro level of global learning theories to micro levels of the individual to the mediated environments that involve the materials, tasks and settings of the learning. Specific language skills are often at the heart of many studies and the assessment of student outcomes provides yet another set of diverse studies. In order to gain an understanding of the influence the use of technology has on second language learning and teaching, it is necessary to review the literature related to the various modes and aspects of technology enhanced instruction in the classroom and at a distance vis a vis second/foreign[1] language oral output. One of the goals of this study is to inform FL instructors about out-of-class and distance education speaking tasks and assessments.

The following discussion focuses on the six most significant aspects of second/foreign language acquisition literature in relation to the basic objectives of the study: speaking theories and models, communicative approaches to teaching with tasks & task design, technology and language teaching/learning, FL speaking offline/online, speaking assessments and the evolution of distance education.

SPEAKING A FOREIGN LANGUAGE: PROCESSES AND HYPOTHESES

Most of what early researchers believed about speaking another language had roots presented by first language theories and models (Chomsky, 1970; Levelt, 1989).

The important question to ask was *how* people learn languages. Chomsky's theory deals with linguistic knowledge, not language usage while Levelt (1989) provides a first language model to spoken production. Krashen (1985) focuses on language input as the source for acquisition while Swain (1985) claims that output plays an important role in acquisition.

Krashen (1981) has several hypotheses about L2 processing. Krashen's way to *acquire* a language is to use it for communication and language *learning* is a conscious endeavor. His *monitor hypothesis* suggests that "learning is only usable as a way of checking things acquired. The *affective filter hypothesis* puts together several individual variables …as a "mental block" or "filter" that can get in the way of acquisition." (Cook, 2001:190) Finally, Krashen's input hypothesis claims L2 learners acquire language by understanding meaningful messages. Only through comprehensible messaging can "humans acquire language" in L2. Krashen's model of L2 use demonstrates the relationship between acquired language and language output. The process of speaking another language depends primarily on acquired knowledge, which is monitored with "conscious" grammatical rules.

[1] This study uses FL and L2 to describe a foreign language context.

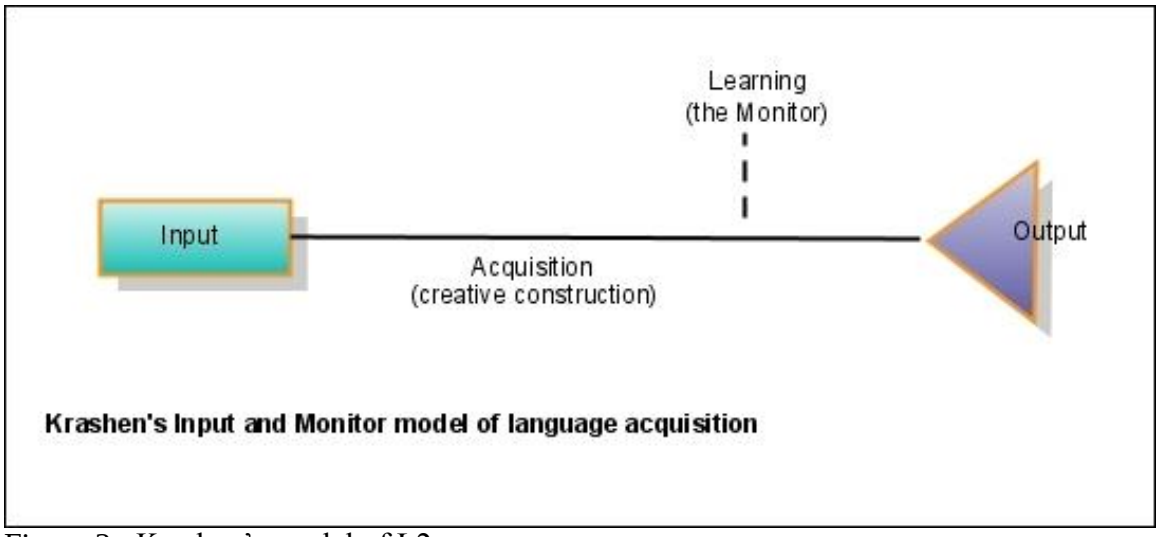


Figure 2: Krashen's model of L2 use

While Krashen argues that input is the greatest determiner of language acquisition, Swain (1985) claims the hypothesis associated with L2 competence is output practice. In Canada, language immersion programs are rich with input learning environments but practice using the target language (TL) is often lacking. Swain's Output Hypothesis places emphasis on L2 learners "noticing" gaps in their linguistic knowledge as a result of external or internal feedback (monitoring). This serves a metalinguistic function of internalizing linguistic form, testing hypotheses about the language and increases control over previously internalized forms (Swain, 1985; 1993).

Levelt's production model is a process with four components: the Conceptualizer, the Formulator, the Articulator, and the Speech Comprehension System (SCS). Utterances begin as preverbal or non-language specific messages in the "Conceptualizer." The Conceptualizer stores the plan in working memory and then the message is sent to the "Formulator" to transform the preverbal message into a linguistic plan. Payne (2002) criticizes Levelt's model for its lack of acknowledging the

importance of short-term storage. It is at the point in the process, short-term storage, that Payne positions for language processing for L2 use.

The Formulator has two processes: to encode the preverbal message to form a syntactic structure and to select the phonological representation. The phonological plan for articulation is stored in the "Articulatory Buffer." This is used by the speaker for internal self-monitoring in the SCS. If there is a match between what is stored in the Articulatory Buffer with the intended message, the speech motor functions produce the utterance as overt speech.

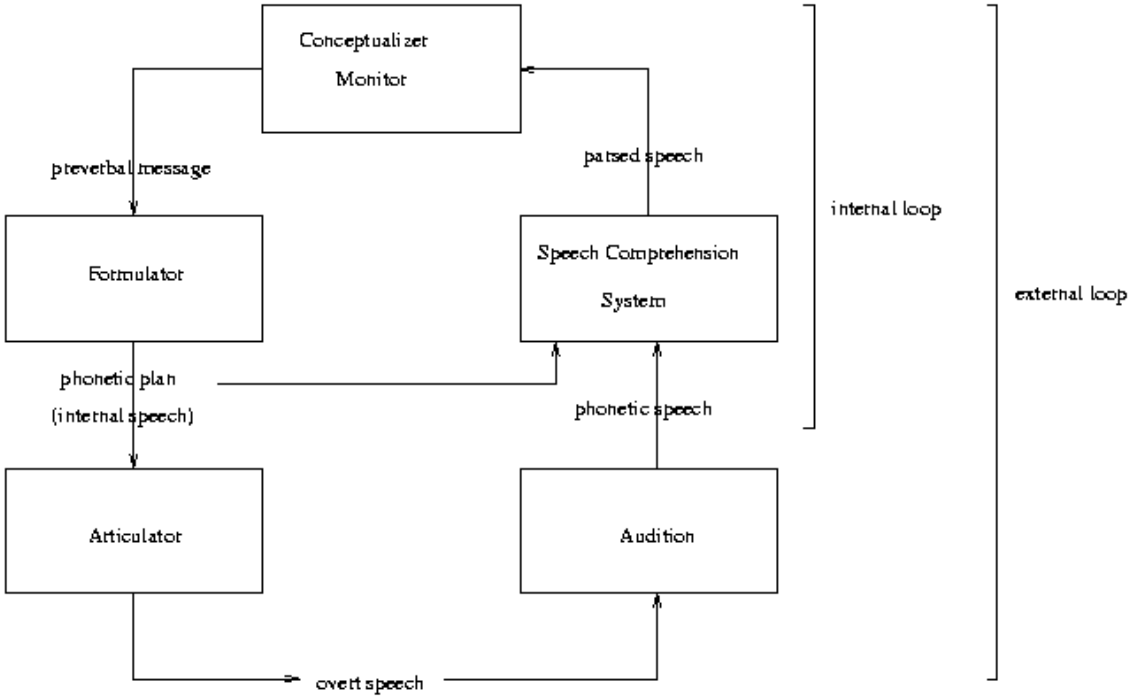


Figure 3: Levelt's Production Model

The linguistic component consists of two subsystems, one for production and one for understanding. This is a double ``perceptual loop" to model that a speaker can attend to her own *internal* speech before it is uttered and can also attend to her self-produced

overt speech. Payne (2002) combines Levelt's Model with working memory theory to guide performance-oriented second language (L2) tasks and suggests that working memory may contribute to individual differences among L2 learners. In the case of speech production in task performance, working memory extracts and stores (temporarily) linguistic and encyclopedic data from long term memory (Ellis & Barkhuizen, 2005). For beginning-level language learners much of this process is dependent on controlled rather than automatic processing (Ahmadian, 2010).

Bachman and Palmer's (1996) model of language ability works with the idea of competence as knowledge and includes strategies. The model's knowledge set includes: organizational, textual, pragmatic, grammatical, and sociolinguistic. The strategies are three-fold: goal setting, assessing, and planning. The strategies deal with micro-decisions the speaker must make in order to perform. Bachman and Palmer suggest that the strategies utilized may be affected by setting and/or by test task. Their implications as elements of task design for this study will be further analyzed in chapter 3 when discussing methodology used in this investigation.

As interesting as the theories and hypotheses are concerning language production processes, they do not have a way to account for varying performances on L2 production tasks. Working memory is involved in everyday performance in L2 use and in the classroom. The restrictions on the learner's ability to speak the language are caused just as much by memory limits as by the difficulty of the syntax or vocabulary (Cook, 2001) or TL input.

COMMUNICATIVE APPROACHES TO TEACHING – TASKS & TASK DESIGN

The communicative approach to teaching begins with the theory of language as communication. Linguist Dell Hymes (1974) uses "communicative competence" as the goal for language teaching. This was in contrast to Chomsky's theory of competence –an idealized version of a speaker in an environmental vacuum. Hymes saw language in the contextualized version of people using language in community (culture). Canale and Swain (1980) take the Hymesian model and suggest four competencies: strategic, grammatical, sociolinguistic and discourse. Strategic competency refers to coping strategies of initiating, terminating, maintaining and repairing communication. Grammatical competence refers to the formal aspects of communication (grammatical and lexical abilitities). Sociolinguistic competence refers to understanding the social context, and discourse competence refers to the interconnectedness of how meaning is represented in relationship to the entire discourse or text. This in turn has been elaborated by Bachman (1991). The Bachman model has been further developed and expanded by Celce-Murcia, Dörnyei, and Thurrell (1997).

At the level of language theory, communicative approaches to teaching have an eclectic, theoretical base. Some of the characteristics for the communicative view of language include:

- Language is a system for expressing meaning
- Language is intended for interaction and communication
- Language structure reflects its functional and communicative uses.
- Language units are categories of functional and communicative meaning as used in discourse.

The Communicative Classroom

Within the frame of communicative approaches to teaching foreign languages is an instructional set of pedagogic interventions called *Focus-on-Form* (FoF) (Ortega, 1999). These types of activities are designed to engage the learner to specific forms of the language in the context of meaningful use. "*Focus-on-form* instruction represents an alternative to both focus on forms and focus on meaning" (Long, 1991: 123). "*Focus on form* entails a focus on formal elements of language whereas focus on form**S** is limited to such a focus, and focus on meaning excludes it" (Doughty & Williams, 1998: 4).

There are three tenets to FoF. First, it presupposes the learner is engaged in the meaning before attention to the code occurs. Second, the choice of form is at the appropriate level for the learner. Lastly, the targeted form(s) is/are primary in the communicative needs of the instructional task. One aspect of the activity is pre-task time for planning. What occurs during planning time is left to the learner who is free to assess task demands and regulate linguistic resources. It is believed that planning may lessen the cognitive load of a task and free up attentional resources at the micro levels of speech production. Furthermore, with pre-task planning, it is claimed that there may be increased opportunities for noticing the gap or holes in one's competence (Schmidt & Froda, 1986; Swain, 1998).

Ellis (1987) was the first to suggest that forms that have not been automatized by the learner are more likely to access by allowing time for planning. Crookes (1998) proposed that planning as a condition of a task pushed output to a learners interlanguage (IL) limits and engages SL acquisition processes. Skehan (1998) suggests that the

opportunity to plan before an L2 task lessens communicative stress which encourages focus on form. Using pre-planning as an element of task design, therefore, can act to balance the competing goals of meaning and accuracy.

Another set of communicative class activities is presented by Rivers and Temperley (1978). They present a framework that bridges receptive and usage skills to prepare students for authentic communication. The processes involved in learning to communicate are in a two part system: skill getting and skill using. They do not represent successive stages since beginners have skills they can use to communicate at least something. However, they report that there is a gap to bridge the two processes to help learners make the transition. They suggest that using *pseudo-communicative skill-getting activities* would lead "naturally" to spontaneous communication (1978: 5). They propose a bridge between both processes and suggest instructors give students activities to help them transition from acquisition to oral production.

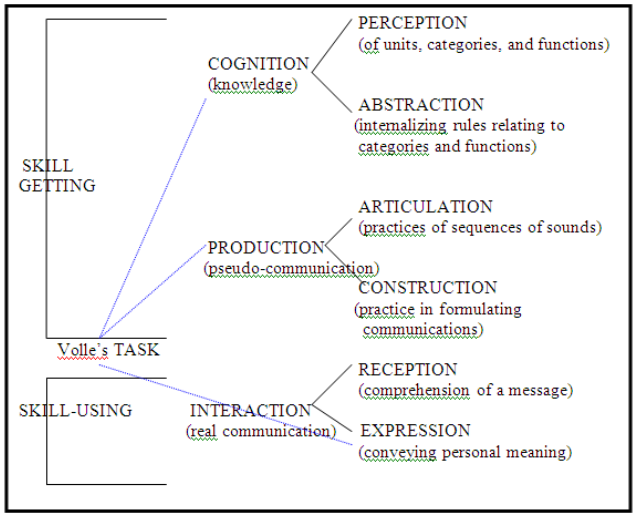


Figure 4: Modification of Rivers and Temperley's Framework for Developing Speaking Skills (1978)

There are two levels or steps of activities. The first one is to prepare the student to "achieve" through structural exercises and quasi-communicative activities. The second activity type includes functional communication and social interaction activities. The Rivers and Temperley framework is important to the current investigation since the "bridge," the speaking task, has an intended listener and allows for personalization of information.

Task-based Instruction (TBI) is an approach to teaching linked to theories of SLA that has a considerable base of research (Ellis, 2003). The features of "task" in SLA are multi-faceted. According to Tarone and Yule (1989), a task provides the students with three components: some pre-selected information to convey, information to complete the task, and the awareness that an information gap exists. They note that different tasks will elicit different amounts of discourse from single sentences to long monologues (124).

Several studies have investigated the effects of types of task whether a "one-way" or "two-way" task type on a learner's performance (Pica et al., 1993; Shortreed, 1993). A two-way task is an exchange with two participants negotiating meaning while a one-way task has one participant giving information to another without feedback (Lee, 2002).

Task Definition

The definition of "task" is wide and varied. Bygate et al. (2001) suggest that any definition of task be supplemented depending on the "different purposes to which the tasks are used" (11). For Ellis (2002), a task is meaning focused language use, which involves a language user not a learner where the focus is both on meaning and form.

For Long (1985), a task is a "piece of work undertaken for oneself or for others, freely or for some reward. Thus, examples of tasks include painting a fence, dressing a child, filling out a form…In other words, by *task* is meant the hundred and one things people do in everyday life, at work, at play and in-between" (89). This was the broadest of all definitions.

Johnson (2003) defines a task as either an "activity or exercise: what we give our students to do in classrooms" (4). For the purposes of this study, a task refers to a pedagogic activity that involves a piece of "classroom" work that engages learners in comprehending, manipulating, producing, or interacting in the TL, based on Nunan's (1989) understanding of this term.

While there are narrow and broad views of tasks, the term "authenticity" also has come to be known as a dimension of a task with its own range of meaning. Taylor (1994) describes four types of authenticity: text-based, learner's interpretation of text, tasks conducive to language learning, and situational/contextual. Contrary to tasks might be meet the L2 learner's comprehension, Nunan (1989) defines authenticity as coming from any material which has not been specifically (re)produced for the purposes of language teaching. "Authenticity" for speaking tasks for this study is one dimension of the task design that should reflect a real-life situation or context (Ros i Solé & Mardomingo, 2004; Skehan, 1998). Authenticity should be recognized "in the act of interpretation" (Widdowson, 1979: 165).

While a *task* has many definitions, this investigation uses the term interchangeably with activity or prompt that link the class objectives and authentic contexts to communicate in the TL.

Task Design

Task design depends on set goals, aims and priorities of the instructor or curriculum. Task design usually reflects the theories of learning (Rosell-Aguilar, 2005). Ellis (2000) argues that task design should incorporate the concepts from second language acquisition (SLA) as processing input and output, but also from a sociocultural perspective. According to Ellis, "the psycholinguistic approach provides information that is of importance to *planning* task-based teaching and learning" whereas the sociocultural approach is concerned with the performance of the task in an improvisation that promotes communicative efficiency and L2 acquisition" (193). Task design should provide a purpose to learners, encourage "interaction," and promote TL use as a means to an end.

There are additional dimensions to task design: planning as it relates to cognitive demand and familiarity of task. The bulk of research (Crookes, 1989; Ellis, 1987; Ortega, (1999); Wigglesworth, 1997) has gone into planning time. The studies hypothesized that planning time would allow L2 learners to produce more accurate, lexically more complex and more fluent speech. The results are mixed, especially in regard to accuracy: while some studies showed more accuracy (Ortega, 1999; Ting, 1996; Wigglesworth, 1997), other studies found this increase was dependent on task (Foster & Skehan, 1996; Mehnert, 1998) or on certain measures (Crookes, 1989). Ortega (1999)

concludes that a firm evaluation of pre-task planning is not possible given different interlanguage[2] measures and the different operationalizations of "planning."

More recently, Ahmadian et al. (2010) distinguishes between "careful" with-in task planning from "pressured" with-in task planning. Careful planning allows speakers to attend carefully to the formulation stage during speech planning and engage in pre-production and post-production monitoring of their speech (Yuan & Ellis, 2003). "Pressured" planning requires the learners to produce language under time pressure (Ellis & Yuan, 2005).

Yuan and Ellis (2003) conducted a single-factor between-participants design with three levels of planning condition (no planning, pre- task planning and on-line planning). This study was done outside of a class context with learners of English but not directly tied to a specific curriculum. It differed in findings from other studies in fluency (Foster and Skehan, 1996; Mehnert, 1998; Wendel, 1977), reporting that pre-task planning leads to more fluent language use. The variable of time may suggest the difference in findings. Yuan and Ellis (2003) had time pressure in the design of their study while the others had no such time pressure.

In addition, research shows that learners' performance improves if they feel in command of the situation, and if they are familiar with their environment (Oxford, 1999; Rosell-Aguilar, 2005). Repetition tasks are viewed as assisting beginning and intermediate L2 learners since their process hinges on controlled rather than automatic processing (Ahmadian et al., 2010). Ellis (2003) suggests that with task repetition "when

[2] Interlanguage first introduced by Selinker (1972) to refer to the language learners produce at any point in time resulting from systematic rules that characterize learner progression. Emphasis is placed on

learners know what they are going to talk or write about they have more processing space available for formulating the language needed to express their ideas with the result that the quantity," fluency, and complexity will be enhanced (246-7).

Task design as situated in communicative approaches to language learning contain crucial elements to promote language output: authenticity, familiarity of task, planning time and task repetition. This next section will discuss technological components to class activities.

Evolution of Technology and Language Learning Approaches

The early use of technology in foreign language learning environments came from and extension of the applied uses of the phonograph and radio and later audio-visual labs (Salaberry, 2001). In the 1960s, new technologies involved programmed instruction based on the audio-lingual method (ALM). This was the combination of structural linguistics and behaviorist psychology that resulted language learning via conditioning (Omaggio, 2001). Two primary speaking practice activities of ALM are dialogue memorization and oral pattern drills. There were two types of oral pattern drills: repetition drills and transformation drills. Repetition drills were repeated after the model while the transformation drills required a minimal change that was reinforced after the tape recorder (Omaggio, 2001). Computer use during this time and into the 70s replicated this orientation to language learning and promoted the behavioristic tutorial-and-test approach, also known as "drill-and-kill" (Hanson-Smith, 2002: 212).

"interlanguage" as a system in its own right as well as it being a stage in development.

In reaction to ALM, the 70s and 80s saw many pedagogical approaches such as Suggestopedia, Silent Way, and Community Language Learning. Computer use at this time began to flip the "computer-as-tutor" and allow for the learner to have control to "do something" within simulated environments (Kern & Warschauer, 2000: 14). These types of computer-based activities correspond to cognitive/constructivist views of learning.

By the early 90s, criticism of the instructional technologies in the foreign language context was strong because the computer-assisted language learning programs were not integrated into the courses effectively, and were only used ad hoc. Another criticism concerned the physical disconnect of the teacher from the student performing the computer-based activity because it could "compromise the collaborative nature of the classroom" (Kern & Warschauer, 2000: 16).

The mid to late-1990s saw computers used by learners more than just a receptive tool to provide input. The main goal is to aid the learner "to develop their own constructions of language" (11). Computers became viewed as mediational tools that link people to each other through a networked system or out into the World Wide Web.

Computer usage and language teaching does not represent a particular technique, method or approach. CALL activities are language learning tasks that teachers design based on the belief that they have a positive impact on student learning. Many reports discussed the advantages of using technology in and outside the classroom as well as online[3] and offline[4]. It is at this juncture CALL blends into the Internet assisted language learning to distance education class sites.

[3] Online refers to the use of the Internet.
[4] Offline refers to technology that is not need the Internet, i.e. CD-ROM, DVD, etc.

TECHNOLOGY IN FOREIGN LANGUAGE LEARNING/TEACHING: SPEAKING ONLINE & OFFLINE

During the mid-1990's personal computers began appearing as a regular part of the American home. However, researchers abroad seemed to have been the forerunners in investigating distance learning with foreign language courses (White, 1995). White's work (1995) was one of the first publish on the ideas of autonomy, strategies, expectations and beliefs of the distance language learner. By 2003, differentiating research lines emerged specifically for languages in distance education: learner variables (Hurd, 2004; Little, 2001; White, 1995, 1999), course development (Fox, 1997; Doughty & Long, 2003) and online language course testimonials with preliminary findings (Cahill & Catanzaro, 1997; Despain, 2003, Gamlin, 2005).

Speaking Offline

Technology can provide for increased opportunities for students to use their own voices to practice in the target language (TL). Ho (2003) conducted a study to employ audio-taped dialogue journals to investigate the effects for EFL learners. This out-of-class set of tasks, similar to written journals, was shared back and forth with the instructor for comments and feedback. Twenty-six intermediate level students participated in a one-semester oral communication class. The oral journal entries were 20% of the final grade. There were no claims to support improvement: "I am not able to quantify the extent to which the students' oral competence improved, because progress is not easy to measure" (276). She evaluated to number of entries and observed student behavior with the assignments. She found that six students did not comply with the

assignments and showed limited language ability. On the other hand, "shy" students found the privacy of recording provided a stress-free environment for speaking-aloud practice. With this type of oral practice, Ho notes, students "seldom need to use such communications strategies as paraphrasing or circumlocution to compensate for a lack of language" (276).

Speaking Online

Hoopingarner (2004) and his team from Michigan State developed an online Audio Portfolio program that allows for students to record their narrations.
This, in essence, replaced old audio cassettes that students recorded onto outside of class or in a language lab. It makes managing the audio recordings accessible for both students and instructor available 24/7 and from anywhere there is an internet connection. The new tools replace the old ones and with advantages: organizing speaking files and providing feedback.

What can technology do as a tool to facilitate language learning? Cahill and Cantanzaro (1997) first introduce on "online" Spanish course for Christopher Newhart University. Learning outcomes were compared between classroom students and online in writing and observations were made about other aspects of the design of the course with materials. The finding for writing showed a clear difference in favor of the on-line students. The reason for the significant difference was "unclear" (109). The speaking aspect of course design was outlined.

There were two online course development phases. The first model did not include any multimedia software. Speaking tasks were completed in steps: (1) written

dialogs between two partners and (2) a ten-minute three-way telephone conference with the partners reading aloud the dialog and the instructor providing feedback to the students. Other speaking activities were individual and submitted by leaving a telephone message to the instructor. The researchers found that first time Spanish students had poor pronunciation and added interactive CD-ROM program to the next phase of the course.

Although the audio cassettes for the lab provided much input for learners, it appeared not to translate to speaking skills. The second phase of the online course included an additional offline component. The inclusion of the interactive CD-ROM program allowed students to listen, record and compare their own speech to a native speaker. The authors report that this improved student pronunciation from the first course. There were no measures to support this claim. The speaking activities with the instructor changed as well. Students no longer created dialogs but were "evaluated" in five minute telephone calls with the instructor. No further description was given about the type or kind of evaluation. It was noted that the speaking practices are not traditionally part of the final course grade.

Kötter, Sheilds and Stevens (1999) were the first to discuss online audio conferencing with text-based tool to investigate oral performances. It was predicted that the participants would develop greater fluency rather than accuracy but a definitive conclusion could not be reached. Emphasis in the report was on the tasks as promoting language development in the target language (TL). A relevant finding was that in spite of students participating anonymously in the synchronous chat, participants were more "reserved" (58) in comparison to written chat; students revealed that they were more aware of mistakes and gaps by participating in this mode.

Despain (2003) presents findings from a 30-month beginning Spanish study in f2f, f2f with TELL, and Internet[5] delivery modes conducted at North Carolina State University. The first semester there were two delivery modes: f2f and f2f with TELL. Subsequent semesters added the Internet-based student sections. Each class session was recorded and streamed as a live Webcast via the Internet and archived for later delivery for Internet-based classes. There were a total of 49 f2f students, 47 f2f with TELL students, and 38 Internet students. Significantly different rates for retention were found. Internet-based students withdrew from the course and much higher rates than f2f and f2f TELL sections. Although there were no differences found for course completers on achievement scores on hourly texts and final exams, very little was reported on the oral aspects of the study.

Internet students were "encouraged" to watch each class session with another student to complete exercises as though they were in the class. Students worked with a partner to complete a series of 21 paired activities (one per class session) and create a report via e-mail. In addition, Internet students "were required to submit three voice samples during the semester, with the intended purpose of providing further accountability for oral production. For the first two submissions, students read specific readings from the textbook. For the third submission, they responded to a series of questions" (247). How these were scored and what part of the final grade these Internet oral activities played were not mentioned. The hourly tests and exams did not include a

[5] Internet refers to the course delivery method. The internet provided course content, but students were encouraged to meet in f2f. One criterion for a Distance Education course is that students are also remote from each other.

speaking component. With the encouragement of Internet-based students to meet in person, no tracking of the f2f meetings were offered.

The use of the Internet has provided students with opportunities to view parts of culture outside their own as well as communicate with native speakers (Hertel 2003). Hertel discusses the use of email exchange of Spanish learners with ESL students in Mexico as having provided motivation to learn more about another culture and use the target language. The motivation or engagement appears similar to the pen-pal programs often facilitated by language instructors throughout the world.

Gamlin (2005) shames past CALL research claiming that sophisticated technology for interactive computer games and software abound, but yet there is a continued "…inability of academic technology to adequately assist learners with speaking a target language" which reflects poorly on the CALL field (19). As a solution, he suggests integrating sound or voice files into the curriculum to enrich both student input and student practice in online environments such as Blackboard[6]. He provides a list of speaking activities/tasks and reports online sound files may be used in an online assessment, which appear more as listening based as opposed to a speaking based.

Volle (2005) explored speaking online with 19 DE students enrolled in a first semester Spanish course. She investigated students' pronunciation in two types of recorded speaking activities and two synchronous conversations. Students created voiced audio e-mails to send each week to the instructor during the semester. These voiced files were read aloud and grammar-based completion drills. In addition, students participated

[6] Blackboard is a course management system often used to provide course information, activities and communication.

in two- Internet mediated oral conversations with their instructor using MSN messaging at the mid-term and final. Three kinds of data were scored and compared from the conversations: articulation, accuracy and proficiency. The findings suggest that only in the area of proficiency were there significant gains in scores. Blake et al. (2008) suggest that her lack of control over several variables "makes the results uninformative" (115).

Blake, Wilson, Cetto and Pardo-Ballester (2008) investigated using a telephone based speech recognition software program to compare the oral language proficiency of f2f, hybrid,[7] and distance learning (DL) formats. The course materials for the hybrid and DL sections included CD-ROM/DVD programs, online content-based web pages, and synchronous chat with text/audio. The f2f sections used a textbook in lieu of the CD-ROM/DVD program. The data generated by the test shows that all students reach comparable levels of oral proficiency during the two years of study.

The program for first year is divided into three levels (1-3) of beginning Spanish. The second year is divided into four levels (4-7) of intermediate Spanish. F2f students met in the classroom five days a week, the hybrid students met in class for two, three-hour classes, and the DL students met online one hour per week with two partners to do collaborative content-based tasks. Measuring Proficiency is based on Levelt's (1989) model of speech production and comprehension. The test is conducted via the telephone and lasts 20 minutes. The exam required testers to read aloud, listen and repeat, say the opposite, answer short questions, build sentences from jumbled-up word combinations,

[7] Hybrid refers to blended learning: partial classroom and computer mediated instruction and practice

answer open-ended questions, and retell stories. The test estimates the tester's level of automaticity[8] with the language.

Although Blake's study was to compare the oral proficiency of his program and not to explore speaking performance gains, one step that would have strengthened the study might have been to include a pre-test for all students. There are study design questions that persist. Because the study expanded across quarters, how many times did students repeat the exam within the numbers reported? Did students change learning delivery modes during the study's length? What kind of feedback did student receive after taking the 20 minute exam? How closely aligned were the tasks to the curriculum?

There are many areas in comparing student oral performances yet to be explored as is noted from the studies above. The interest in this exploration must be situated within a proper L2 developmental context.

SPEAKING ASSESSMENTS: PROFICIENCY, ACHIEVEMENT, AND PROCHIEVEMENT

The general purpose of a language assessment is to determine someone's *proficiency* in a language. The problem is how to define the construct of proficiency – a point that has been debated for quite some time.

ACTFL's Oral Proficiency came about from two main avenues: the government and the universities. WWII opened the eyes of the military and U.S. agencies to the fact that most Americans were not multilingual. Post WWII, the Foreign Service Institute (FSI) was interested in improving its employees speaking skills. In 1952, the agency also began to develop a valid speaking test. In 1956, the FSI had their first oral proficiency

[8] Automaticity is the ability to access and retrieve words, create phrases and clauses and give responses

interview test, and it became mandatory for all State Department employees (Lowe, 1983). Over time, the American Oral Proficiency Interview Test became the testing instrument for the CIA, FBI and the Defense Language Institute.

At the time of development, the test was seen as a predictor to communicate and "not as a direct measure of speaking ability" (Sollenberger, cited in Clamphan & Corson, 1998: 76).

The Peace Corps worked with the U.S. Department of Education (ETS) to create materials for teaching and learning to include FSI-style tests. In the mid-70s, DLI, ETS and Peace Corps combined to form the Interagency Language Roundtable (ILR). By the late 70s, then President Carter called for a report to determine the status of foreign languages in the U.S. colleges and universities. One of the recommendations that came from the report was the establishment of standards for L2 testing (Fulcher, cited in Clamphan & Corson, 1998).

It is at this juncture that ACTFL, The American Modern Language Association (MLA) and ETS began to work on procedures of standards for both government and academic use. The standards were based on the ILR. Initially the scales did not capture the lower speaking levels of beginners so that zero, one and two levels of two were included in a scale to five (Lowe, 1983).

Proficiency in second language theory has been debated. In the 1950s, in the proficiency movement, it meant the same as *achievement* (Lowe, 1988).

---

without conscious attention to the linguistic code (Cutler, 2003; Levelt, 1989).

The foreign language teaching profession is currently attached to the term "proficiency" in terms of describing one's place in the L2 using continuum. "Proficiency" as evaluated from the Oral Proficiency Interview (OPI) is not sensitive enough to capture progress with in first year or one semester language courses (Blake, 2008). To maximize the washback[9] effect in testing, it is important to design testing criteria that reflect the course objectives and rubrics for rating the task, knowledge or skill the learner should be able to demonstrate.

The following figure demonstrates the washback effect for students, teachers and researchers (Bailey, 1998: 264). The test directly informs the student that may support or change the student learning strategies or processes. This in turn will support the next testing session which become a cyclical process. Teachers are also affected by tests. The teacher can change the materials and/or the instruction and/or the exam itself to achieve course objectives.

[9] Washback Effect refers to the extent the test effects student learning and instructor teaching.

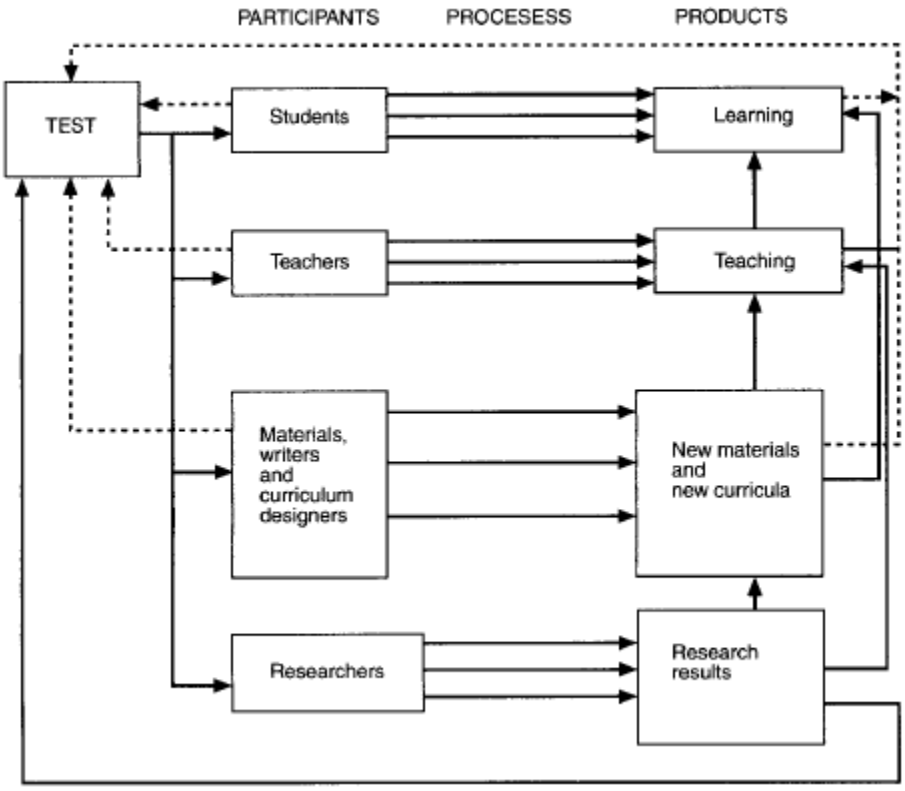


Figure 5: A Basic Model of the Washback Effect

Although the above is part of an entire washback program, it is also appropriate for single test to a series of semester tests. Bailey (1992) points out that there can be a clash when students have two goals: an intermediate goal of getting a good grade on a test and a long term goal of improving proficiency. The test can, therefore, be seen as having a negative washback to the learner. To avoid this, an appropriate connection between developmental language level (from within the curriculum) and test items need to be made.

The present study is interested in investigating student oral abilities as achievement measures of course goals and to determine if the creation of speaking tasks tied to the curriculum assist students reaching course or personal goals. With this focus,

the use of theoretical constructs of communicative competence and proficiency shows that it is necessary to introduce another concept, that of "prochievement." The literature cites Gonzalez Pino (1989) and Hendrickson (1992) coining the term to indicate a hybrid form between proficiency and achievement test formats while still focusing on assessing students' progress in the "classroom."

Shrum and Glisan (2000), define prochievement testing as the "blending of achievement and proficiency testing that incorporates linguistic (structural, syntactic, lexical) items from the course content and objectives to test achievement but also elicits language performance to assess ability to function along a proficiency-base continuum" (296). For this reason the speaking tasks prompts used in the present study were designed with prochievement in mind. Their suitability for the goals of this study will be further analyzed in chapter 3 when discussing the methodology used in this investigation.

STUDENT PERCEPTIONS OF THE CALL EXPERIENCE

There are just a few studies that have asked student to describe and rate the quality or their experience in online language courses: Adair-Hauck, Willingham McLain and Earnest-Youngs, 1999; Chenoweth and Murday, 2003; and Green and Earnest-Youngs, 2001.

Adair-Hauck, Willingham McLain and Earnest-Youngs (1999) used a self-report questionnaire to compare students' opinions in their blended (hybrid) CALL French class with students in the f2f class. They found more students in the hybrid section reported meeting their personal language learning goals than those in the f2f section. The students in the hybrid section noted that they perceived being able to spend more time on difficult activities because the technology allowed 24/7 access. The students also felt less stress

working alone rather than with a f2f group. There seemed to be more learner control and student-centered learning.

Both Green Earnest-Youngs (2001) and Chenoweth and Murday (2003) administered surveys to online and offline students. Both studies report less than favorable experiences online. Students from both the CALL and f2f sections reported the same satisfaction with their semester's language achievement. Those who completed the CALL section found that some web page material was too difficult and the coordination of activities was not well organized. The online students in Chenoweth and Murday's (2003) French study had a lower overall course rating than f2f course ratings.

Murray (1999) also reported on the student experience using CALL materials. He interviewed students who used an interactive videodisc to study French for one semester and received responses similar to Adair-Hauck, Willingham-McLain, and Earnest-Youngs (1999). In Murray's responses, students reported that they liked working at their own pace and spending time on difficult activities. Self-directed learning is again highlighted.

While no research has reported specifically on students' perceptions and beliefs of CALL materials for oral development or CALL speaking performance confidence, Krashen's theory (1981) suggests that self-confidence is specifically identified as an important aspect of the "affective filter" in that it enables the learner to encourage intake - (useful) input.

While it is well documented that severe levels of anxiety works to debilitate foreign language performance, small amounts may help to facilitate performance. Students who choose to take distance education courses have full schedules – between family and work obligations. Taking on additional academic responsibilities may increase stress levels. The nature of distance learning – learners isolated in space and

time - may have built in it, a specific anxiety. To date, there is only one article that concerns itself specifically with anxiety and in distance learning (Pichette, 2009).

Pichette (2009) sought to discover if there were different anxiety profiles between first semester language classroom and distance education students. The study included 186 French speaking students studying English and Spanish. The study reported two findings. There were no significant differences in anxiety levels between: in class and on line foreign language learners or between the language classes (French and Spanish). The study suggested that students with more experience with the language had less anxiety compared to those without experience (82). Pichette's study suggests that the lack of anxiety difference between Canadian students and U.S. students is, perhaps, due to the fact that there is prior L2 experience (85) in Quebec.

There is a clear lack of research regarding the perceptions of students about using CALL in distance education for the development of oral language skills.

CONCLUSION

Our "academic" world is fast changing in terms of where we learn and technologies we use to support or facilitate that learning. Twigg (cited in Russell, 1999) states that "No matter how it is produced, delivered, whether or not it is interactive, low-tech or high-tech, students learn equally well with each technology and learning as well as their on-campus, face-to-face counterparts even though students would rather be on campus with the instructor if that were a real choice." With the publication of "No Significant Difference," Russell illustrates with over 200 research articles that the findings suggest no differences in terms of learning outcomes and coursework achievement.

The studies included in Russell's report span from 1928-1998. Included in these studies reporting no significant difference were eight language studies. The first five studies looked at the effectiveness of the "audio-Laboratory" in elementary language courses and found that there were no significant differences listening to the teacher versus to the tapes in the lab (Brushwood & Polmantier, 1953; Fotos, 1955; Buka et al., 1962; Lorge, 1965 & Sisson, 1970). These five studies have the teacher removed from the students while working in the language lab and therefore, fulfill the requirement of "distance."

Lou, Bernand and Abrami (2002) conducted a meta-analysis of undergraduate distance education based on media supporting the DE pedagogy. Of these 103 studies, spanning from 1985-2002, there were two of the studies included foreign languages. One study compared web-based versus classroom learning for a course on teaching ESOL and found that there was no end-of-course difference in achievement among the students enrolled. However, because the students in the online section began with statistically different pretest scores, the authors suggest that the students in the online section made greater overall gains than the offline students (Thirunarayanan & Perez-Prado, 2002).

The second foreign language study detailed a second year Japanese course using a televised class format versus a traditional class format without finding a significant difference in course outcomes. (Kataoka, 1987) Although this course design is classified as distant from instructor and supports the findings that suggest that DE students learn equally well at remote or host sites, it fails to inform us about the skill of speaking or speaking performances. Neither study suggests that speaking performance was isolated

and compared. The basis of comparison was course competence with paper and pencil rather than performance.

A broader meta-analysis about teaching courses online summarized the research based on course environment, learner's outcomes, learner's characteristics and institutional and administrative factors (Tallent-Runnels et al., 2006). This report contained 76 studies from 1995 to 2004 that included only two studies about foreign languages. In this case both were about FL teacher education.

Because students mostly report that their primary objective is to "speak" or "communicate" in the foreign language for immediate, intermediate or long-term goals, oral development in formalized classes continues to be a concern for the FL profession. With today's rapidly changing definition of "formalized" classes, the ability of L2 speaking needs to be investigated to provide the best practices for promoting stated goals. The current study of speaking assessment scores under two conditions of required versus optional speaking and DE versus f2f can lead to the tasks and context that contribute to FL oral development. The methodology for the present study is outlined in the next chapter.

Chapter 3: Method

INTRODUCTION

Because of the paucity of research investigating speaking skills framed in distance education, this quasi-experimental study seeks to examine the oral performances of first-semester Spanish students in two learning delivery modes: in an online, distance education (DE) context and in a face to face (f2f) class context. More specifically, this study seeks to compare performances among students at the beginning, middle and end of the semester of study to determine the overall quality of the oral performances with required speaking activities versus the oral performances without required speaking activities. The study uses elicited information from surveys and post-exam self-assessments to describe the students' perspectives in developing their oral skills as well as any insight into their actual use of materials.

The four cells below represent the characteristics of the classes and can be seen in Figure 1.

Four 1st Semester Spanish Course Contexts

	Mode of Delivery	
attribute of oral task	Distance Education Section with required speaking practice	In-class Section with required speaking practice
	Distance Education Section without required speaking practice (optional)	In-class Section without required speaking practice (optional)

THE RESEARCH QUESTIONS

The specific research questions are as follows:

Question 1: Are there differences among the sections' performances, regardless of delivery mode and assigned speaking tasks, on the same speaking tasks recorded at two points in a semester?

Question 2: If there is a difference between students' performance as a result of the different delivery modes, what is the nature of the difference?

Question 3: How do the different class sections compare on the final speaking exam when *new*[10] testing prompts are introduced?

Question 4: What do students report on their final exam self-assessment in terms of their own performance confidence levels and beliefs about learning activities? Are there any differences in learner experiences among the delivery modes and those with different amounts of oral activity?

THE RESEARCH DESIGN

This chapter describes the methodology used to collect and analyze the data for this study of oral performances by first semester Spanish students. The following section describes the institutional setting, the participants, the course content, the "classroom" environments, and the speaking tasks. The final section of the chapter describes the instrumentation needed to answer each question, the procedure to collect the data, and the method of analysis.

[10] "New" means not having performed the specific speaking prompt during a previous exam time.

The Institutional Setting

Founded over 40 years ago, Central Texas College is a public, open-admission community college offering associate degrees and certificate programs in academic, professional and vocational/technical fields. With its main campus in Killeen, TX and over 100 other locations around the world, Central Texas College serves over 50,000 students on military installations, in correctional facilities, in embassies and on ships at sea. It offers the first two years of academic course work to the communities of military installations and specific military contracts.

Along with the traditional face-to-face classes it offers, Central Texas College is currently a member of GoArmyU (formerly eArmyU), a consortium of higher education institutions that offer distance education coursework. The courses are offered in three modes: (1) online through a Blackboard management system, (2) a stand- alone CD-ROM course, or (3) PDA course. Each distance education course is required to be "equivalent" to its traditional, in class course. Fall 2003 was the first semester that a foreign language course was offered in the distance education mode of delivery (Volle, 2005).

The Participants

The Students

Students at this institution tend to be active and non-active military personnel and their dependents. The students are located at U.S. Army bases in the U.S. and abroad. (See Map in Appendices) Each course section begins the semester with 27 students. The

students from both f2f sections are located in Central Texas area. The students in the DE sections are located in various places in Texas, the U.S. and abroad.

Students registered for a traditional[11] Spanish class or for an online Spanish class. The traditional classes met either Monday/Wednesdays from 9:00 am to 11:50 am or Tuesday/Thursdays from 9:00 am to 11:50 am. The selected section for required or optional speaking was done at random and students were not aware of the assignments ahead of time. The distance education sections were also randomly assigned to required or optional speaking. Students were not aware in advance of the speaking requirements.

Student gender and language learning background for the fall 2006 first semester Spanish classes are in the Table 1. There are more females completing the face to face sections than males, whereas there are about the same amount of males and females completing the semester.

It is interesting to note, but not uncommon in the Central Texas area, that there were family groups registered together to take Spanish. The Monday/Wednesday section had two sets of siblings: a brother with two sisters and two sisters. The Tuesday/Thursday section had a mother-daughter pair.

[11] "Traditional" is used here to describe a face to face class with six contact hours per week. It seems that the contact hours remain the same as when the "traditional class" completed all tasks in the classroom. Only extended writing was done outside of the class. The f2d sections in this study have the traditional contact hours but in fact, employ many CALL elements.

First Semester Spanish Classes *N*=64	Learning Background 0=No Previous Language Experience 1 or 2 or 3= # of years of high school Spanish 4= Non-English 1st Language 5=Grew up with Exposure to Spanish	Gender
F2F with Required Speaking **n=24**	0 = 5 (20.8%) 1= 4 (16.7%) 2= 11 (45.8%) 3= 1 (4.2%) 4= 1 (4.2%) 5 = 2 (8.3%)	M = 4 (16.7%) F = 20 (83.3%)
F2F without Required Speaking **n=19**	0 = 3 (15.8%) 1 = 1 (5.3%) 2 = 8 (42.1%) 3 = 4 (21.1%) 4 = 0 5 = 3 (15.8%)	M = 6 (31.6%) F = 13 (68.4%)
DE with Required Speaking **n=13**	0 = 3 (23.0%) 1 = 5 (38.5%) 2 = 3 (23.0%) 3 = 1 (7.7%) 4 = 0 5 = 1 (7.7%)	M = 6 (46.2%) F = 7 (53.8%)
DE without Required Speaking **n=8**	0 = 1 (12.5%) 1 = 3 (37.5%) 2 = 2 (25.0%) 3 = 1 (12.5%) 4 = 0 5 = 1 (12.5%)	M = 4 (50%) F = 4 (50%)

Table 1: Frequency Counts for Language Learning and Gender for Fall Spanish Sections 2006

The Instructor/Researcher

I was the Instructor/Researcher, and I am a female non-native speaker of Spanish. I lived, studied and/or worked in Spain, France and Honduras for extended periods of four to 24 months. I am a native English speaker with abilities in French and Arabic. I have been a full-time faculty member, early on in ESL and currently in Spanish and anthropology at the institution for 23 years. My first experience with distance learning was in 1996 with V-TEL. This telecommunication system with two-way audio/ video

combined two classes from different campuses in real time. In 1997, I developed a distance education online anthropology course. In 2001, I began planning and designing online distance education sections of Spanish. I entered into the PhD program in Foreign Language Education in 2002. At the time of the current study, I was the sole instructor for foreign languages online at this institution.

I have witnessed the evolution of *expected* class activities and tasks over the years. What has changed? Early language teaching classrooms were the primary focus for oral practice and feedback. Outside the class, students were expected to study-review of grammar and vocabulary and complete reading tasks (see Figure 6).

Inside the Classroom	**Outside Class**
Grammar explanations by Instructor Oral input Pair and group work Discussion – Instructor lead Feedback by Instructor Authentic written materials Video Audio	Study grammar and vocabulary Practice reading Practice extended writing

Figure 6: Early Classroom Activities and Tasks, circa 1980

With the introduction of CDs and basic Internet, class activities and roles of teacher and student changed. With portable course materials such as CDs that play on any computer, grammar videos (video teacher – learner), interactive language games (computer-learner), audio and authentic online texts shift where the learning/practice take place and the responsibility of learning to the student (see Figure 7).

In the Classroom	Outside Class
Oral input Pair and group work Discussion Feedback by Instructor Instructor facilitates in class activities	Study grammar and vocabulary Practice reading Grammar explanations online Authentic written materials Practice extended writing Video Audio

Figure 7: Language Learning Activities shift, mid to late 1990s

In 2002, online voice tools change where the "class" is physically located. They make it possible for all activities to shift from any physical space. Synchronous versus asynchronous activities, teacher and student presence and (teacher) feedback and (student) self-monitoring are central design features in planning distance education courses.

The Speaking Exam Proctors

There were two oral testing sessions: the midterm and the final. This required the assistance of proctors for the DE students. There were 12 individual test proctors and one primary testing center with additional testing centers located on other campuses. DE Students located in locations without a testing center selected a supervisor-volunteer to proctor the speaking midterm and final exam sessions. The 12 test proctors were located in Alaska, Georgia, Kentucky, Wisconsin, Texas, Colombia, Iraq, and Macedonia. The primary testing center was on the Central Texas College campus in Killeen, Texas.

Each proctor was identified by the student. The instructor corresponded directly with the proctor to verify the acceptance of the testing responsibilities, dates of exams and general procedures for the exams. The test packets were sent directly to the proctors via email attachments. The students were not allowed to ask the proctor questions during

the exam. The proctors returned the testing packets by fax, by mail or scanned in an email attachment. The students were responsible for dropping their voice files into the Blackboard drop box during the testing session. The f2f students were proctored by their instructor in scheduled small groups.

The Speaking Exam Raters

There were two native speaker raters who evaluated all students' oral performances. Both have taught Spanish at the university level for more than three years. The male rater is originally from Ecuador and has resided in the United States for more than 4 years. The female rater is multilingual, originally from Mexico and Italy and has resided in the United States for more than 12 years.

The Transcript Coders

There were two transcript coders. Both were female, non-native speakers of Spanish that have taught elementary Spanish for more than 10 years at the community college level in Texas. Both have studied in Spain and traveled throughout Mexico.

The Curriculum

The Spanish program content was identical for each section. It emphasized a multi-skilled approach to teaching Spanish and draws from communicative theory and practice from the *National Standards for Foreign Language Learning* (2006). The program integrates the 5Cs of Communication, Culture, Connections, Community, and Comparisons. It offered many opportunities to connect, compare, and communicate with local and virtual Spanish-speaking communities. In terms of speaking skills, this course

focused on the interpersonal mode of communication (ACTFL's Communication, Standard 1.1; 484).

As the core component of the program, the textbook was divided in half for two semesters of study with the preliminary pages through chapter seven for the first semester and the continuation of the textbook for the second semester. This study was concerned with the first semester of study.

A typical chapter is comprised of seven sections in the following sequence:

1. Opener: Communication Goals & Cultural Context(s) of Communication
2. Vocabulary: mini-video dialog & lexical audio
3. Video & U.S. Community Voices
4. Grammar & Cartoon
5. Cultural Explorations
6. Reading
7. Writing

The chapter's opening two pages presents the communication goals and cultural objectives for the chapter. It provides cultural information and asks students to make guesses about information from the Spanish-speaking countries to be presented later in the chapter.

The vocabulary sections present active vocabulary in photos and in mini-video dialogs that are sub-parts of the chapter's longer, video dialog. The themed vocabulary is practiced through activities that progress from structured to more open-ended. Each chapter has two to three cycles of vocabulary presentation and practice.

The video viewing section presents a new video viewing strategy with each chapter. There are pre, during and post activities to test comprehension. Each chapter includes an extension of the theme to a cultural presentation with extra practice of the

cultural material. This cultural presentation features a prominent Hispanic living in the United States. This also offers students opportunities to explore other sources of information on the use of Spanish in the professional fields.

The grammar sections include both presentation and practice. There are three or four cycles of presentation and practice moving from formation to use. Authentic materials (Nunan's definition of realia) are used to see the "grammar" in use in a realistic context. The activities in this section progress from controlled through communicative practice. This section is also divided in the middle by the presentation of a cartoon. In the earlier chapters, the cartoon tests students' comprehension and beginning in Chapter 5, the activity expands beyond to ask students to express their reaction and go beyond the basic content of the cartoon.

The cultural exploration section focuses on the chapter's featured country or countries. The tasks students perform required that they make guesses about meaning and use strategies to improve reading and cultural comprehension. Cultural readings activities are modified so that the students do not need to understand every word of the cultural information in order to complete them. The tasks are organizational and analytical in nature. The theme also ties back to the chapter opener.

The reading selections contain modern topics that are taken from authentic texts. The questions are simplified for easier comprehension and completion. While the reading sections are strategy based, the writing sections are process based. Strategies assist the students in approaches to the text and comprehension of the reading. Process based writing presents a series of steps that build the writing content into more polished paragraphs or short essays.

The writing tasks build sequentially across the chapters from audience and topic selection through writing a topic sentence through paragraph development. There are pre-

writing, writing and revision steps. Often, the details come from the chapter practice with other students. The variety of written pieces includes e-mails, reviews, descriptions, letters, journal entries and narrative stories.

The following table compares the type of course components and learning tools associated with the DE and f2f sections.

Course Component	F2f Sections	DE Sections
Online Class Components	Nicenet: Announcements, Documents, Discussion Board, URL links	Blackboard: Full Course Management System Eduspace: Content & Wimba live classroom
Textbook: Nexos Media Edition, 2007	Physical Textbook with audio CD (Nexos)	Eduspace: Publisher eBook (Nexos) w/ Wimba
Workbook/Lab Manual	Workbook for outside class activities Lab Manual: Instructor plays audio in class	Eduspace: e-Workbook e-Lab Manual
Dialog & Culture Videos	Instructor plays in class	Student plays in eBook
Interactive CD-ROM	Student uses outside class	Student plays in eBook
Grammar Videos	Instructor plays in class	Student plays in eBook
Companion Website - ACE the Test Activities - Improve Your Grade	Student links in Nicenet	Student links in eBook

Table 2: First Semester Course Components for f2f and DE Sections

The f2f students accessed materials differently from the DE students. The DE students had regular access to an eBook located in the Eduspace, a publisher provided course management system. Vocabulary and grammar presentation are embedded directly on the pages of the eBook. The students control when and how often to play the vocabulary audio files and view the linked grammar video tutorial lessons directly from the eBook pages (see Appendix H: Chapter Lesson Plan).

Illustration 1 below exemplifies the "interactive" nature of the eBook pages. The arrow reveals the icon with the clickable movie. The megaphone icon indicates that individual audio vocabulary files are linked beneath each word on the page. When the online student is learning the vocabulary, she can click on each word or selected words and in any order. She can repeat the word or listen as many times as she feels necessary. The f2f students did not have access to an eBook, but could access vocabulary by playing the audio flashcards on the companion website.

Illustration 1: A Sample Page of the Interactive eBook

For the DE students, the activities from the interactive CD-ROM and the companion web site are embedded within the eBook. The eBook contains matching, drill, listening, reading, activities that provides immediate feedback while other activities that are more open-ended in nature had to be sent to the instructor. The companion website or

online learning center included activities in two categories: Improve Your Grade and ACE the Test.

The *Improve Your Grade* component is interwoven throughout the text to provide online audio flashcard practice for vocabulary and grammar concepts, the in-text audio of vocabulary, dialogs and narrations, and expanded in-text activities to web searches and specific web pages to review and support themes of connecting to the community, culture and reading.

The ACE practice tests were available for vocabulary, grammar and video comprehension. The activities were automatically scored for immediate student feedback and not linked to a grade book. Students submitted paper or digital reports with their scores and questions for review.

The syllabus was similar for all four classes except for grade distribution. One section of each delivery mode (f2f & DE) was assigned graded speaking tasks for each chapter. These three to four speaking tasks per chapter were due on the same dates and were graded with detailed feedback. These speaking tasks were worth a total of 20% of the course final grade.

Those sections not receiving graded speaking tasks had access to them as optional speaking opportunities. Students who submitted any of the optional speaking tasks received the instructor's detailed feedback but no points were awarded toward the final grade.

The course syllabus presented students with chapter objectives and schedule for a chapter of study with specific chapter quiz and exam dates. DE Students received a "study guide" providing a sequence of activities to complete over a 10-day period, due

once a week or twice per chapter. F2f students received this guidance in class because their homework was due four times per chapter.

The time frame provided to complete one chapter in the f2f class was two weeks or 12 hours of class meetings. The time frame to complete one chapter in the DE class ranged between 17 and 20 consecutive days. The following table illustrates the first semester pace.

First Semester Spanish	**Week 1**	**Week 2**	**Week 3**	**Week 4**	**Week 5**
	Orientation	Ch 1	Ch 1	Ch 2	Ch 2
	Week 6	**Week 7**	**Week 8**	**Week 9**	**Week 10**
	Ch 3	Ch 3	Ch 4-**E**xam	Ch 4	Ch 5
	Week 11	**Week 12**	**Week 13**	**Week 14**	**Week 15**
	Ch 5	Ch 6	Ch 6	Ch 7	Ch 7-Exam

Table 8: The Pacing of First Semester Spanish

The DE sections received the syllabus in Blackboard, the required course management system for the institution. The f2f classes received hardcopies of the syllabus and had access to an electronic copy on an "unofficial" course website called Nicenet (www.nicenet.org). This was set up by the instructor.

The "Classroom" Setting

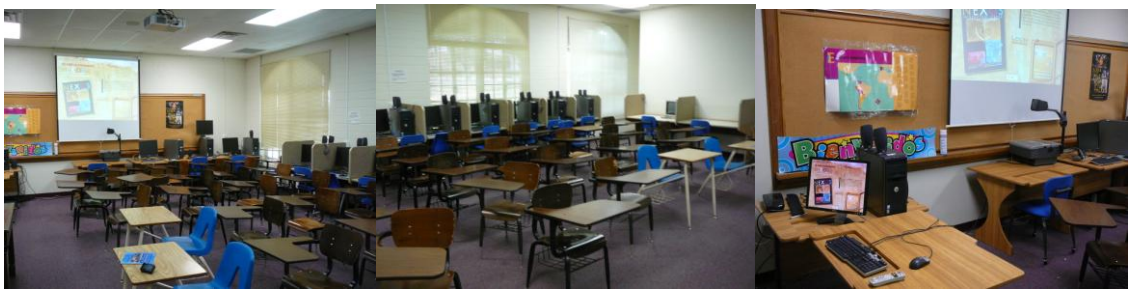
The f2f Classroom

The classroom for the f2f sections not only had the physical space, but also a virtual space, Nicenet, for out of class communication (see illustration). Nicenet provided the f2f class an online discussion board for all class and small group discussion boards for asynchronous written communication. It also contained an announcement

section that was used to add the homework instructions. Having a class webpage was especially helpful for students who missed class or students who left class quickly at the end to reach another part of campus as they could access the announcements to print out the homework instructions. The announcements saved time and helped lower students' stress levels for missed work. The class web pages also contained various URL links that supplemented the course topics. The final feature helped students stay connected outside of class with email. By accessing the student roster, an email link allowed student to send messages that were archived as sent and received in the system and copied to the personal email account.

Illustration 2: The f2f Online Class Website

The f2f classroom contained ten multimedia computers with Internet access. Students could use these computers before and after class. The computers were also available on Fridays. For special activities, students worked in teams in class collecting information or posting group findings after discussions. There was a main computer with projection to a screen. The computer played DVDs and MP3 files and displayed Internet sites. There was also a TV with a DVD/video cassette player. The chairs were movable to create either small-group circles or traditional rows. The instructor was in the room and set-up the activities and learning tasks for each scheduled meeting.



Illustrations 3, 4 and 5: The f2f Classroom

The DE Classroom

The DE classrooms were primarily through two course management systems: Blackboard and Eduspace. Blackboard is one of many systems commonly used in higher education to maintain grades, course content and communication. It is the required system of this study's institution (see illustration below).

Illustration 6: The DE Blackboard Course for First Semester Spanish

Eduspace contained the e-materials for the DE students (see Illustration 7). Eduspace is a Blackboard branded course management system for the publisher of the textbook. In order to have the majority of course content and communication consistent, Eduspace was the chosen location for written discussion boards and the Wimba live classroom sessions located via the "communication" tab.

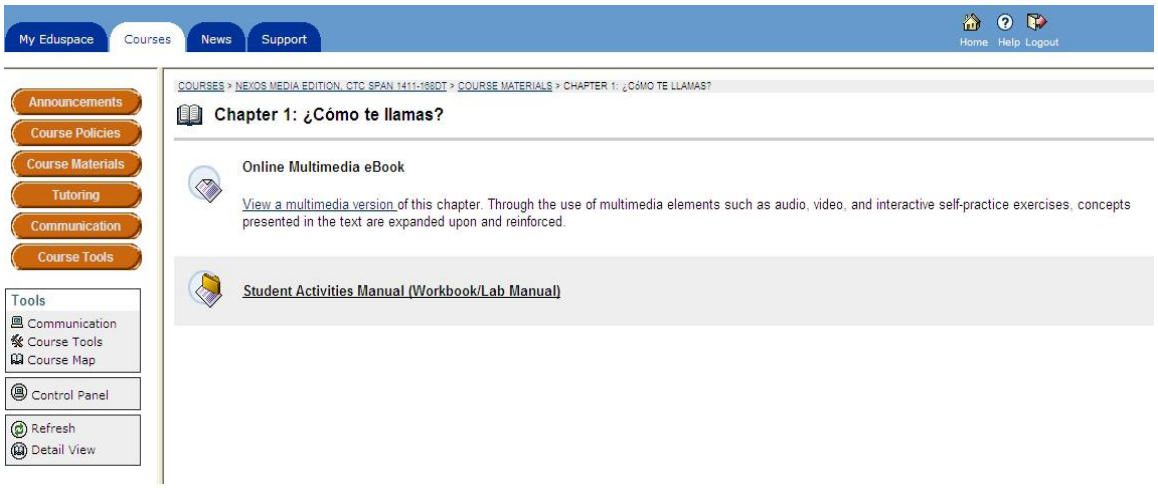


Illustration 7: The Eduspace View of Links to eBook and eWorkbook/Lab Manual

Embedded in the eBook speaking activities are prompts to the student to open Wimba live chat rooms. As students work through practice exercises and tasks, they meet classmates virtually to complete the activities together in real-time; seeing the activity and talking and/or texting to their partners at the same time. The following illustrations provided screen shots of the classroom. The left screen shot shows the participant view and the right screen shot demonstrates instructor view.

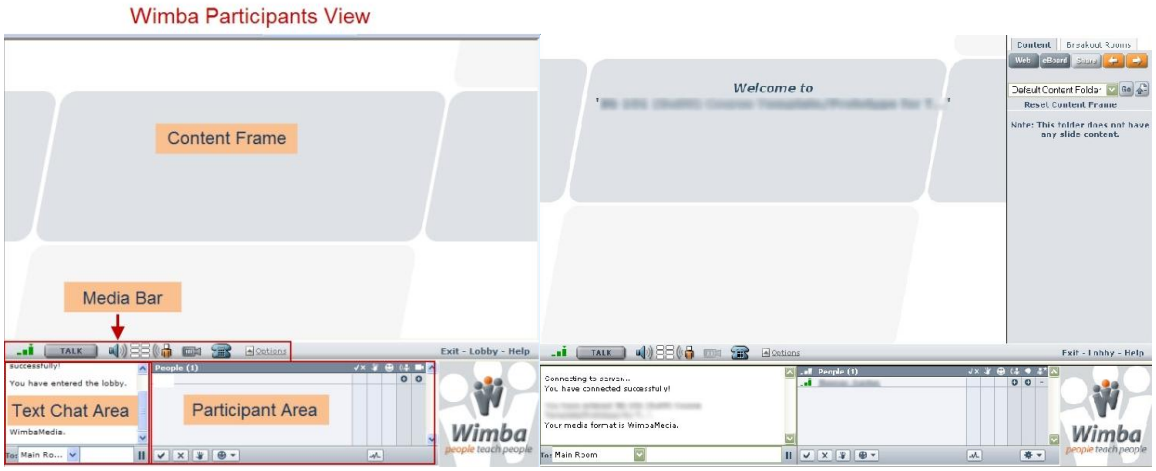


Illustration 8: Wimba Live Classroom Participant & Instructor Views

Both DE sections had one scheduled live class meeting per chapter. These lasted 30 – 45 minutes per session and were archived for later viewing/listening. This time was

used to answer questions regarding activity instructions, review concepts or provide extra practice as well as introduce the next chapter's topics. Students who did not participate in the live classroom were "required" to access the class session archive.

Wimba live class was included in the price of the student course bundle as part of Eduspace (e-materials & live classroom for $5.00 per student). In addition, DE students had a paper textbook and interactive CD-ROM as a backup in case of connectivity or technological issues; especially important for those students located in remote areas of the world.

The f2f students purchased a textbook/student activities manual & CD-ROM bundle for about the same net price of $90.00. For local DE and f2f students, there were bookstore issues in the purchase of the correct bundle. One bundle had the Eduspace code and no physical student activity manual (workbook/lab manual) and the other bundle contained the student activity manual and did not have the Eduspace code. This caused a delay for several students at the beginning of the course until an exchange could be made at the bookstore.

The Speaking Activities, Required/Optional Speaking Tasks, and Exams

The Textbook

All students were required to participate in the textbook speaking activities that were assigned. DE students worked with each other inside one of the ten Wimba classrooms. Each Wimba live classroom was named for a country to be studied during the first semester. F2f student worked with each other during class and outside of class as needed. Each chapter had several speaking activities for each subsection, but not all activities were assigned. Here are examples of assigned speaking activities.

> Introductions are a normal part of everyday life. Study the drawing and, with a partner, create four short conversations in which one person introduces another person to a third party. (Long, 2006: 14)
>
> Introduce yourself to another classmate. Exchange information about where you live, phone numbers, and e-mail addresses. (20)

In the DE classroom, student work was somewhat monitored. It was impossible to provide feedback to every student. DE students often set up weekly "team" online hours to do the oral practice together in small groups. These were not recorded or monitored.

Here are examples of speaking activities that were not assigned.

> With a classmate, role-play a cell phone conversation in which one of you has dialed the wrong number. You are curious about the person you have accidentally reached. Try to get as much information from each other as possible. (12)
>
> You are looking for something you need for class. Tell a classmate what you're looking for, and he or she will tell you if that item is or is not in the classroom. If what you ask for is there, he or she will tell you where to find it. (24)

Neither of these speaking prompts was chosen to perform in class or online. They appeared too contrived. The DE students were not in a traditional classroom environment. Therefore, the second prompt would need to be rewritten for their learning context.

Required/Optional Speaking

There were five sets of speaking tasks distributed throughout the semester. Each set corresponded to themes and structures from chapters 1, 2, 3, 4, and 5/6. All students in all fours sections were provided with the speaking tasks at the beginning of the chapter. This allowed students to make mental note of an item that might support the speaking task.

At any time in the lesson, the student could begin to produce his electronic speaking file set. Each speaking task within the set could be produced with a microphone and any PC's basic record function. This produced a WAV file. Because of the size of the file, f2f students copied their files to CDs or added them to their USB storage device. DE students uploaded their files from their computers into the Blackboard drop box.

Here is an example of a chapter speaking task set.

Capítulo 3 – Speaking Task Set

A.  B. 

CH03 01WAV file

In the above drawings say what these people ARE doing using the verb directly and where they are doing it. Start with "En dibujo A, ... En dibujo B....."

CH03 02 WAV file

Using the same A & B drawings, describe what they are going to do on Sunday.

CH03 03 WAV file – Mi horario Personal

Tell me about your classes this semester. How many classes do you have? What are they? When are they? Are they easy (fáciles) or difficult (difíciles)? Tell me if you work and the schedule you (generally) keep each week. (Los lunes etc….)

CH03 04 WAV File

Tell me your name, date and current time (your time).

Illustration 9: Example of Chapter Three's Speaking Task Set

The speaking task sets included both simple sentence level tasks as well as more personal and open-ended speaking. In this example set, WAV file 3 was written to involve both an academic schedule as well as a work schedule to reflect the students' real life experience. As much as possible, scaffolding within the task instructions were provided with visuals and written prompts.

Not all sections were required to complete these speaking files. One section of f2f and one section of DE were required to complete the speaking task sets as part of their semester grades. The other two sections were provided with the same speaking tasks sets, but not required to submit them for a grade.

All students who submitted speaking files were provided with detailed feedback. Below is an example of the evaluation sheet used to note vocabulary use (correct versus incorrect lexical item and pronunciation), structural use and overall comprehensibility. The instructor included positive comments as well as instructive feedback for errors.

Ch 03 Speaking Student	WAV 01 A. The family is walking in the park. B. Ana is reading at home.	WAV 02 Going to…. A. B.	WAV 03 Mi horario Personal School/work	WAV 04 Tell me your name, date and current time.
Vocabulary Structure				
Comprehensibility Yes/No/Partly				
Comments:				

Table 3: Example of Feedback Form

Speaking Assessments: Orientation, Midterm and Final - Description and Procedures

The speaking prompts were intentionally designed to fulfill the characteristics associated with a communicative approach to classroom learning, namely student engaging, possible real-life situated prompts. In other words, the speaking prompts might be used outside of class with speakers of Spanish. In addition to authenticity as a main feature, the prompts reflect the topics and themes practiced in the semester with more than one possible appropriate response with the focus on the interpersonal mode of communication.

Another design feature to the midterm and final assessments was the element of a specific time in which to complete the files. The midterm assessment allowed the student 30 minutes to complete both speaking files. This gave students plenty of planning time as well as opportunities to re-do performances. The final assessment also was to be completed in 30 minutes but there were six files to create. Time pressure was evident as the planning time was three times shorter per speaking prompt. Students were in control

to decide which order to complete the files. Therefore, the time devoted to each speaking prompt might have varied with more time spent on one prompt versus another. Much like the Levelt's model, time forced automaticity within the framework of beginning FL students (see Table 4: The Semester's Time Features).

Orientation Speaking Sample	Midterm Speaking Exam	Final Speaking Exam
No Time Pressure Two Speaking Prompts & Technology Steps	30 minutes Two Speaking Prompts	30 minutes Six Speaking Prompts

Table 4: The Semester's Time Features

At three points in the semester the students' speaking abilities were assessed: at the beginning of the course, the midterm and the final. During the orientation week for all sections, students were provided step by step instructions to make WAV files with prompts to complete two initial files. F2f students were given instructions and time in class to produce the two files. The instructor/researcher collected the files from the classroom computers. For DE students, the first initial speaking files were completed without a proctor during the orientation week.

The midterm contained two speaking prompts similar to activities found in the course content. The instructor/researcher served as proctor for all f2f student sessions, reading the same instructions and monitoring the time as set forth for the testing sites and individual proctors throughout the world.

One class session was long enough to test an entire f2f section in groups of eight students in 30 minute intervals. DE students were given a due date to complete the midterm speaking test. DE Students at a testing center or proctored individually had the testing protocol package sent in advance to the proctor/site. The instructor/researcher

communicated directly with the testing supervisor or proctor via phone and/or email. Whether from online via the drop box inside Blackboard sections or from the class computers, the instructor/researcher collected all voiced WAV files on a USB portable storage device for immediate grading and feedback to the student as well as for later rater analysis for this study for course completers.

The final speaking assessment contained six speaking prompts (see Appendix F). Four of these prompts were repeated. The two initial speaking tasks and the two midterm speaking tasks were repeated on the final assessment. There were two new speaking prompts to reflect the last chapters' topics (see Table 5).

Initial Assessment (Orientation Week)	Midterm Assessment (Week 8: 30 min.)	Final Assessment (Week 14-15: 45 min.)
Speaking Task 1 Speaking Task 2		Speaking Task 1 Speaking Task 2
	Speaking Task 1	Speaking Task 3
		New Speaking Task 4
	Speaking Task 2	Speaking Task 5
		New Speaking Task 6

Table 5: Semester Speaking Assessment Items

During week 15 of the semester, the students from f2f class sections, eight students per 35-45 minute testing session, completed the final speaking assessment on the classroom computers. The testing session included the technology check prior to beginning the 30 minute time to create the speaking files. The teacher/researcher served as proctor for the testing sessions. At the end of the testing session, the teacher/researcher collected all speaking WAV files from the desktop to a portable storage device.

During weeks 14-15, the students completing the course online completed their final speaking assessment at their assigned testing center or with their individual proctor. The on-site proctor handed the student the technology check sheet first, and when ready, the student received the speaking test booklet. The proctors returned the materials to the teacher/researcher via fax and upon confirmation of receipt, destroyed the test copy. There was a 30 minute time limit for each student once the student received the speaking test booklet. All speaking responses were recorded on the computer in WAV files and sent to the digital drop box in the Blackboard course.

The Final Speaking Exam – Task Variation and Design

Tarone (1998) claims that a change in topic will change the discourse produced. Although there are mixed findings regarding task variation and L2 oral performances, it is important to look at the features of the speaking task prompts. This next section will introduce the prompts and identify the features based on Bachman's (1990) organizational and pragmatic competence and Fulcher and Márquez Reiter's (2003) interpretive features (instructions and visual cues).

The first final speaking item was introduced in the orientation week. It was repeated on the final exam (see below).

Speaking Assessment Item 1

You happen to run into your Spanish Instructor who does not speak any English. In Spanish, greet her and introduce her to a family member.

This prompt is very basic, but it is used often when out in public places or attending a party or event. The vocabulary is used to greet and identify relationships. The structure is basic subject to verb construction. There are, however, at least two decisions to make: (1) the sociolinguistic address, tú versus usted, which would change the indirect object pronoun to *te* or *le*, and (2) identify the family member to introduce. There is no visual to help with the introduction.

The second final speaking item was introduced in the orientation week. It was repeated in the final exam (see below).

Speaking Assessment Item 2 A Mexican friend is learning about U.S. customs. In Spanish, describe how you (alone or with family or friends) usually spend the weekends. Also include what you are planning to do this upcoming weekend.

What you do during your free time in your city is often a popular topic when getting to know someone new to an area. The vocabulary needed are action verbs, friends, family and weekend (or specific days). The structures needed are present tense, subject – verb agreement and the near future usage of ***ir a*** + infinitive. There are a series of decisions to make: (1) subject(s) to discuss, (2) places and/or activities, and (3) cohesive phrases – *generalmente, usualmente, este fin de semana, este or el sábado*, etc. There are no visuals.

The third item on the final speaking exam was introduced on the midterm (see below).

Speaking Assessment Item 3

One of your Dominican friends really wants you to practice speaking Spanish more often. As you meet for coffee, you discuss your families in Spanish. Choose one family member to describe in detail. Include the following: name, age, physical description, where he/she lives, employment status, and something about his/her personality.

Spanish speaking friends and in-laws encourage language development. Prompt three is an example of a real situation that is often described by students. The vocabulary needed to respond to this prompt are family labels, possessive adjectives (mi, su), numbers, adjectives of description, and verbs (tener, ser, vivir, trabajar). The structure necessary to complete the prompt includes subject to verb form for irregular verbs and –ar and –ir ending verbs. The student needs to be able to use idiomatic expressions to express age (tener años) and know that personality and physical descriptions mostly correspond with the verb ser. This prompt is much more complex than the surface indicates. The student has two main decisions to make: (1) choosing a subject to discuss and (2) presenting the information in a cohesive manner. The visual provides no support to the student.

The fourth prompt is presented as a new item on the final exam (see below).

Speaking Assessment Item 4

You are hosting a high school exchange student from Ecuador for one week. She has never been in the U.S. before and does not speak much English. You need to use your Spanish to explain her week's schedule to her. Today is Sunday. Use the following schedule. Try to make at least 5 sentences.

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Arrive and meet host families	School 8 AM – 3:30 PM 	School 8 AM – 3:30 PM	School 8 AM – 3:30 PM	7:00 AM  Travel to the Capital 	School 8 AM – 3:30 PM	Return Home 
 		Movies 7:00 PM 		Visit History Museum  Return 5:30 PM	Football Game 6:30 PM  Dance 9:00 PM 	Airport 1:00 PM 

There are several opportunities in the Central Texas area to host international students who arrive and spend two weeks to one school year with families. The vocabulary for this prompt contains time phrases and activities. The structure is present tense or near future. The sociolinguistic knowledge to express is register when "speaking" to the "international student." The instructions indicate a minimum of five spoken sentences which should be presented in a cohesive manner. With a five sentence minimum, the test taker should not become overwhelmed with more than five

opportunities in the visual. The visual is packed with times, places and activities from which to choose.

Exam prompt five was introduced on the midterm (see below).

Speaking Assessment Item 5

Using the following pictures, describe what everyone is doing and say whether or not you like to do the same thing.

a)

b)

c)

d)

This prompt is rather open-ended. Each visual has more than one corresponding interpretation. For example, drawing A corresponds to at least four activities presented in the semester: trabajar en linea, navegar la red (el internet), enviar mensajes por correo electrónico, and leer páginas web. The vocabulary for this prompt contains verbs of activity. The structure choice is either subject to verb (present tense) or estar + present participle (present progressive tense) and the use of *gustar* which follows a different syntactic order. Gustar can be used structurally or memorized as an idiomatic expression

"Me gusta (+ verb infinitive form) también." Success on this prompt depends on a logical interpretation of the visuals.

Final speaking prompt six is a new (time one) item. This is one that was not previously presented neither on the midterm nor the orientation (see below).

Speaking Assessment Item 6

You have entered into a Yahoo Spanish chat room. There are many topics crisscrossing the conversation page. Sometimes the questions and comments appear quite random. Here are several questions directed to you. Since you have a microphone, you have decided it is faster to voice your responses in Spanish instead of type your responses.

a) Chica_amiga asks you to describe how you feel after 4 hours of sleep.

b) Miguel_el_guapo wants to know how you are before an important exam or work deadline.

c) Catracho_cuatro asks how you would feel after losing your backpack with your car keys and wallet.

This prompt reflects the anonymous chat room experience. Not only did students report that they have been in chat rooms, they were encouraged to step into Spanish chat rooms (safely) and report back findings to the f2f class or in DE discussion boards. The vocabulary to respond to the prompt includes adjectives of emotion, and there is more than one adjective that adequately corresponds to each situation. There are two structural choices subject-verb (estar) or reflexive pronoun – verb (sentarse). Another decision to make is speaking directly to "Chica_amiga" or speaking in more general terms. There was no visual to interpret.

The following table summarizes the features of the final speaking assessment by item according to organizational competence and task interpretation (see Table 6).

Organizational Competence				Task Interpretation	
Grammatical Knowledge		Textual Knowledge	Sociolinguistic	Prompt Instructions	Immediacy Visual
Vocabulary Syntax		Cohesion/Rhetorical	Register Expressions		
Item One	Relationship Present Tense	Yes	Tú/ Usted	Yes	No
Item Two	Activities Weekends Present Tense/ Near Future Tense	Yes	Yo perspective	Yes	No
Item Three	Family Adjectives Present Tense	Yes	3rd person	Yes	No
Item Four	Schedule Present Tense/ Near Future Tense	Yes	Tú/ Usted	Yes	Yes
Item Five	Activities Present Tense/ Present Progressive	No	3rd person singular/plural	Yes	Yes
Item Six	Adjectives Reflexive in Present Tense	No	Yo perspective	Yes	No

Table 6: Areas of Language Knowledge based on Bachman (1990) and Task Features Fulcher & Márquez Reiter (2003)

The speaking exam was designed to elicit several levels of TL knowledge based on the semester's study and practice. The students must also interpret elements on the exam; namely the instructions and visuals (if presented). *Grammatical Knowledge* describes how individual utterances or sentences are organized: vocabulary and syntax. *Textual knowledge* helps form the type of speech elicited: listing, conversational with rhetorical or conversational organization. *Sociolinguistic knowledge* provides perspective and register. It can also make use of idiomatic expressions. There are many sub-processes performed by the student in order to communicate successfully.

INSTRUMENTATION, DATA COLLECTION PROCEDURES, AND ANALYSIS

Learner Background

During orientation week, all students completed three assignments: (1) a language learning profile, (2) a technology and personal profile, and (3) the two initial speaking WAV files.

The language learning profile was adapted from the European Learning Portfolio (2000). This document allows students, learning or having learned a language formally or informally, to record and reflect on their language learning and cultural experiences. This was adapted for the course not only to help the student reflect on previous experience, but to identify skills and objectives as a guide for the learner in planning and assessing progress (see Appendix A). The goals and objectives selected for this profile were based on those presented in the semester's course materials. The skills represented were coded on the form for speaking (S), reading (R), listening (L), and writing (W).

The instructor/researcher kept the original profiles and made copies for the students to retain during the semester so the students could maintain their "check list" for objectives during the course and update skills.

At the completion of the course, the instructor/researcher analyzed the course completers' profiles to create a frequency distribution for each class section to determine the initial background in formal language learning experience and personal objectives for the course.

The second form, the Technology and Personal Profile, helped to determine the students' outside class environment in terms of school, work and family obligations as

well as to determine the students' technology skills to work with the course materials and general online experience (see Appendix B). At the completion of the course, the instructor/researcher tallied the student technology and personal profiles to create a frequency distribution for each class section to determine the students' initial technology capabilities and experiences. The personal information was also tallied and a frequency distribution was created for categories of life's obligations.

Analysis of Speaking Procedures

Research Question 1: Are there differences among the sections' performances regardless of course delivery and required speaking task sets?

The speaking files for the semester were organized on a USB storage device into folders by section and by speaking file. Speaking files from students who withdrew from the course and/or the study were removed into a "non-study" folder. Participants in the study were assigned numbers and the files were renamed to represent the participant. Before any analysis could begin, two raters were identified and trained.

The training consisted of two half day sessions of four hours each. The first hour consisted of an introduction to the study, the curriculum, and the six speaking task prompts from the final. We discussed the various responses students may have made given the openness and ambiguity of several prompts. The second hour consisted of a discussion about the concepts of comprehensibility, fluency, vocabulary, and accuracy. A general rater form was distributed with the score of zero to five for each of the concepts. The concept point column contained a brief description that corresponded to the point award (see Appendix C). There were six "non-study" audio files used to

calibrate both raters and the researcher. The scoring was discussed after each rater scored the audio.

The next two hours were devoted to scoring the audio files. The raters did not know if the audio file was from orientation, midterm or final. All files related to one prompt were played in sets. After each set of audio files were played, a short break was taken. At the end of the first day, only two sets of audio files were scored.

The second day the following four audio sets were played and scored. There were two sets of audios that were much shorter to score because the participants had only one file submission. In all, four prompt sets contained 128 audio files and two prompts contained had 64 audio files.

In order to conduct a statistical analysis of each file, the statistical package SPSS was selected. The researcher entered the data from the rater sheets and conducted a Pearson's product-moment to establish the coefficient (Pearson *r*) for measuring inter-rater reliability. The two-raters positively correlated ($r = .891$). For this correlation, a significance level of .05 was adopted.

One of the limitations identified in previous research was the sensitivity of small scales with small cells (Volle, 2005). Therefore, a composite score was reached for each speaking file in order to compare overall performances. This combined all sub-scores into an overall performance score. The minimum score was zero and the maximum score was 20.

Using SPSS, a two-way ANOVA was run to compare four performance sets between the DE and f2f sections which further divided the groups by required speaking or optional speaking. This created four cell class groupings and compared the mean

scores at two points in time. This procedure was chosen over four single *t*-test analyses in order to reduce the chance of a Type 1 Error. In addition, a Chi-Square analysis for independence will be performed.

In choosing nonparametric tests of significance, Chi-Square is probably the most frequently used by educational researchers in causal-comparative studies. A Chi Square (X^2) statistic is used to investigate whether distributions of categorical variables differ from one another.

Research Question 2: If there is a difference between students' performance as a result of the different delivery modes, what is the nature of the difference?

When a difference in speaking performances was identified, the audio files for that speaking prompt were transcribed by the researcher. Locating two coders and providing code training followed. The coders were individually trained and independently coded the interlanguage measures for lexical range and morphological accuracy. Interrater reliability measures performed by SPSS, yielded a Pearson-product moment coefficient of high reliability with $r = .94$.

The "utterance" was chosen as the unit of analysis (Crookes, 1990). Because these were short narratives, the entire oral file constituted the unit to analyze. As a measure of lexical range, two measures were identified: lexical density (words per utterance) and lexical complexity (type-token ratio). Lexical density was a simple count of words per utterance. A word is defined as spoken target language meaningful unit. In this case, "ja, ja" might be considered one word unit to express laughter in Spanish. In contrast, the use of "um" as filler, would not be considered a word. Other exclusions are stuttering, hesitations, non-target language words or repeated syllables (Riggenbach,

1991). Lexical complexity was calculated by the different word types divided by the total number of words (Crookes, 1989; Ortega, 1995, 1999).

Morphological accuracy was measured by errors in inflectional use of the target language. In Spanish, this includes inflected words such as articles, possessives, adjectives, quantifiers, subject number/gender, verb agreement, and contractions (Pica, 1983; Ortega, 1999).

The coded data was entered into SPSS and an analysis of variance was performed on the three interlanguage measures in order to determine if there were significant areas of difference.

Research Question 3: How do the different class sections compare on the final speaking exam when *new* testing prompts are introduced?

There were two speaking prompts that were new to the participants on the final speaking test: test items four and six. Using SPSS, an analysis of variance was used to determine if there were significant differences among the groups' performances.

Research Question 4: What do students report on their final exam self-assessment in terms of their own performance confidence levels and beliefs about learning activities? Are there any differences in learner experiences among the delivery modes and those with different amounts of oral activity?

At the completion of the final speaking assessment, all students were given a self-assessment/reflection questionnaire to complete (see Appendix D). The students were able to view the title of the prompt and provide a self-confidence rating based their performance on a scale one through four. Four was the highest confidence level and one was the lowest confidence level. Immediately following each confidence rating, students were provided an area for comment.

Why measure confidence[12]? The data from the confidence ratings were ranked into high (4, 3) and low (2, 1) and compared by class and speaking prompt. The open comments regarding performance on the six prompts were categorized and provide qualitative data regarding cognitive load, lexical or grammatical accuracy, and anxiety.

The last two questions on the final self-assessment questionnaire were open-ended and intended to elicit course activities that the students believe were helpful in developing speaking skills during the semester. The identifications were categorized and counted for frequency within each class and across classes.

CONCLUSION

This chapter described the methodology used in this study. Participants included 64 students: 21 in a distance education introductory Spanish course and 43 in a face to face classroom based introductory Spanish course. This study collected data from two orientation profile forms, two orientation speaking files (time 1), two midterm speaking files (time 1), and four final speaking files (time 2), two new final speaking files and one final self-assessment/reflection questionnaire. This chapter described the methodologies used for collecting and analyzing data. The next chapter presents the results of the study.

[12] Confidence is often referred to as self-efficacy: A belief in one's ability to do things that they try to do (Pajares, 2000). This is exploratory in this study where autonomy and independence of the student is essential to success.

Chapter 4: The Results of the Study

INTRODUCTION

This chapter first presents the results from the quantitative data, beginning with basic descriptive statistics of the data set and then the results for each research question. This section is followed by a discussion of the qualitative data (i.e., the open-ended questions on the final self-assessment questionnaire.)

Descriptive Findings

The orientation questionnaire was distributed the first week of the fall 2006 semester to four classes of first semester beginning Spanish at Central Texas College. The questions concerned: technology background, current environment, previous language background and personal goals for taking the course. A total of 64 students completed the course whose data I was able to collect for this study.

Technology Background

Students were asked to provide information about their computer hardware, software and ISP (internet service provider). All students (100%) had internet capabilities, a working email address, and basic software knowledge (i.e., word, power point). All but 1 student had a working microphone and speakers. The one student without a microphone was from the f2f optional speaking section.

Students were asked to provide information about their experience online with discussion boards, online chat and online classes. The DE sections both had 100% experience with online discussion boards while 70.8% (n=17) from the f2f required section and 36.8% (n=7) from the f2f optional section did. Online chat was much more comparable among the sections. 91.6% (n=22) f2f required section reported using chats regularly while 89.5% (n=17) for the f2f optional section, 62.5% (n=5) for the DE optional section and 76.9 (n=10) of the DE required section reported using online chat. Both of the DE sections reported 100% previous experience with an online class while both the f2f sections had 50% experience with online classes.

This suggests that the students enrolled in all four sections of beginning Spanish in this study had a high degree of technical experience and facility with the computer.

Student Lifestyle & Environment

This section of the questionnaire elicited information regarding the number of classes the student was enrolled this semester, the number of work hours outside the home, the number of family members this person was caring for and if the participant planned a move during the semester. This personal information was collected in order to provide of full picture of influences students face today with competing attention to family, work, and other coursework.

The number of classes that students reported taking ranged from one to two classes, three classes or four to five classes. Of all the participants, 20.3% (n=13) were enrolled in one or two classes while 32.8% (n=21) had three classes and 46.9% (n=30) had four or five classes. Not only did more than half take more than two classes, most of

these students were employed outside the home. Competing course work may influence the dedication to time for studying.

The DE required speaking section had 100% (n=13) of full-time employment. The DE optional speaking section had one student not working outside the home, one student (12.5%) working 25 hours, one student working 30 hours and five students (62.5%) working full-time. The f2f sections were split between working and not working outside the home: f2f required 50% (n=12) and f2f optional 47.4% (n=9). Employment pressures might influence available time to devote to study.

In terms of living with family members, only two of the 64 students were single students living alone or in the dorm. In the DE sections, eight students were deployed soldiers while six were wives of deployed service men. The f2f sections had family members enrolled together with them in class. There was a three-sibling and two-sibling group in the f2f required speaking class. The f2f optional class had a mother and daughter enrolled.

There were six students that reported plans to move during the semester: three from the f2f sections and three from the DE required speaking section. In addition, one of the soldiers was based in the U.S. and traveled to Colombia on short missions throughout the semester. With disruptions of routine, this might influence focus and timely assignment completion.

Previous Language Learning Experience

Based on the responses from the Initial Student Learning Profile obtained during the orientation of the course, the responses were broken down into six categories of

previous language learning for 64 student participating in the study: No previous language learning (13 students or 20%), one year of high school Spanish (10 students or 15.6%), two years of high school Spanish (27 students or 42%), three years of high school Spanish (6 students or 9.3%), L1 is Korean (2 students or 3.1%) and Spanish is/was spoken at home (6 students or 9.3%). The data is found on Table 7 and summarizes the student previous language experience by class.

		Previous Language Experience						
Class		No previous Language Experience	1 year of HS	2 years of HS	3 years of HS	Non-English As L1*	Spanish At home	**Total**
f2f with required speaking tasks	**Count**	5	4	11	1	1	2	**24**
	%	20.8%	16.7%	45.8%	4.2%	4.2%	8.4%	**37.5%**
f2f w/out required speaking tasks	**Count**	3	1	8	4	0	3	**19**
	%	15.8%	5.3%	42.1%	21%	0%	15.8%	**29.7%**
DE w/out required speaking tasks	**Count**	2	0	5	0	1	0	**8**
	%	25%	0%	62.5%	0%	12.5%	0%	**12.5%**
DE with required speaking tasks	**Count**	3	5	3	1	0	1	**13**
	%	18.8%	38.5%	18.8%	7.7%	0%	7.7%	**20.3%**
Total	**Count**	**13 (20%)**	**10 (15.6%)**	**27(42%)**	**6 (9.3%)**	**2(3.1%)**	**6 (9.3%)**	**64**
	%	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

*The first language of both students is Korean. Killeen, Texas has the largest population of Koreans outside Soul, Korea.

Table 7: Previous Language Learning Experience

The table can be interpreted by looking at the frequency counts in the categories for each class beginning with row one and reading left to right. The final right column represents the total count of students with the percent out of all students. The final row represents the total count of all students belonging in that category regardless of class

section. Most students (of all 64) have at least two years of pervious high school Spanish. Normally this is equivalent to one semester of university level Spanish. The cases in each class of previous language learning produce no significant differences across the study populations.

Student Top Objectives by Skill for the Semester

From the Initial Student Learning Profile, students identified language tasks that they would most like to be able to do. This also corresponded to the open-ended question on the modified European Language Portfolio Survey. From the responses, language skills could be identified and categorized.

The majority of each section identified speaking as a semester skill they would most like to be able to perform. Table 8, The Semester Language Objectives, summarizes the skills and class percentages.

Class		Self-Reported Top Objectives Categorized by Skill		
		Comprehend Spanish	**Speak or Orally Communicate in Spanish**	**Other: Vocab. or Writing**
f2f with required speaking tasks	**Count**	**2**	**20**	**2**
	%	**8.3%**	**83.3%**	**8.3%**
f2f w/out required speaking tasks	**Count**	**5**	**13**	**1**
	%	**26.3%**	**68.4%**	**5.3%**
DE w/out required speaking tasks	**Count**	**1**	**7**	**0**
	%	**12.5%**	**77.5%**	**0%**
DE with required speaking tasks	**Count**	**1**	**11**	**1**
	%	**11.1%**	**84.6%**	**25%**
Total N=64	**Count**	**9**	**51**	**4**
	%	**14.1%**	**79.6%**	**6.2%**

Table 8: Semester Top Language Objectives

The majority of each section ranging from 68% to 85% selected "communication or speaking" as their primary skill objective for the semester. Ho (2003) notes that students that were motivated for improving speaking sills created more oral submissions. If this is the case, one would expect students enrolled in the optional speaking sections to produce WAV files. Contrary to Ho's findings, only two students (7.4%) from the optional speaking sections chose to do those speaking activities throughout the semester.

Summary of Student Background

The majority of students enrolled in this semester's beginning Spanish sections are non-traditional students with families, work, and other classes that compete for the students' time and energy. They have experience with technology and the Internet as well as two or more semesters of formal Spanish. Most students reported that their primary objective, in terms of skill development and abilities, was to speak (or orally communicate) in Spanish.

STUDENT SPEAKING SAMPLES

The following two speaking samples in the form of WAV files were taken or provided during orientation week. In order to determine if students began the semester with significantly higher scores, the samples were scored by two independent raters. Later, the scores were analyzed using a one-way between-groups analysis of variance with post-hoc tests.

Orientation Speaking Results Sample Item One

Students were registered into f2f and DE sections of first semester Spanish (randomly assigned required speaking). Based an analysis of variance, there was no significant difference at the $p<.05$ level in Sample One scores for the four groups [$F(3, 60)=.12$, $p=.94$]. The actual difference in mean scores between the groups was small. The effect size, calculated using eta squared, was .006. Post-hoc comparisons using the Tukey HSD test indicated that the mean score did not differ significantly for f2f with required speaking ($M= 8.08$, $SD=5.24$), f2f with optional speaking ($M= 8.95$, $SD=5.78$), DE with optional speaking ($M= 7.88$, $SD=5.85$), and DE with required speaking ($M= 8.54$, $SD=4.05$). Figure 9, Comparison of Sample 1 shows the comparative means.

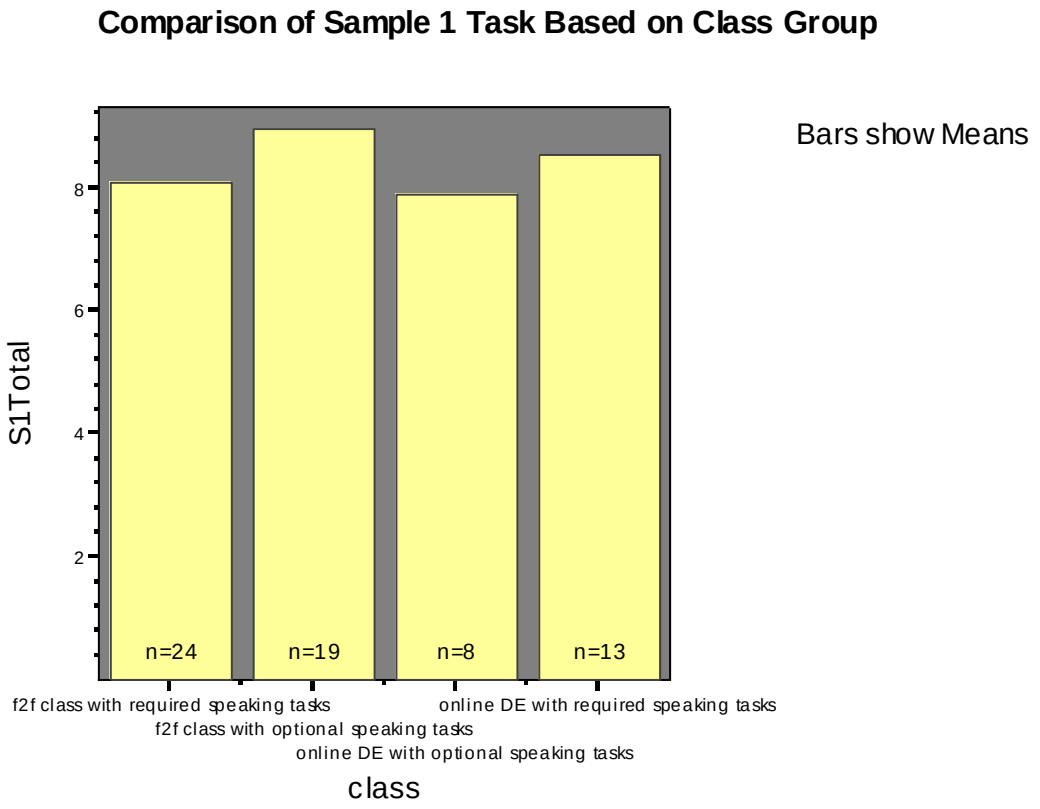


Figure 9: Comparison of Sample One Mean Scores

There was no significant difference among the four groups. Each group began this study with similar speaking abilities.

Orientation Speaking Results Sample Item Two

The results of the one-way between-groups analysis of variance with post-hoc tests are as follows:

There was no significant difference at the $p<.05$ level in Sample Two scores for the four groups [$F(3, 60)=1.1$, $p=.33$]. The actual differences in mean scores among the

groups were small. The effect size, calculated using eta squared, was .05 which in Cohen's (1988) terms would be a medium effect size.

The post-hoc comparisons using the Tukey Honestly Significant Difference (HSD) test, the most appropriate statistical measure to employ when no a priori hypothesis exist with respect to groupings, indicated that the mean score did not differ significantly for f2f with required speaking (*M*= 1.13, *SD*=3.24), f2f with optional speaking (*M*= 1.74, *SD*=3.26), DE with optional speaking (*M*= .5, *SD*=1.41), and DE with required speaking (*M*= 0, *SD*=0). Figure 10, Comparison of Sample Two, reveals the initial performance means among the groups.

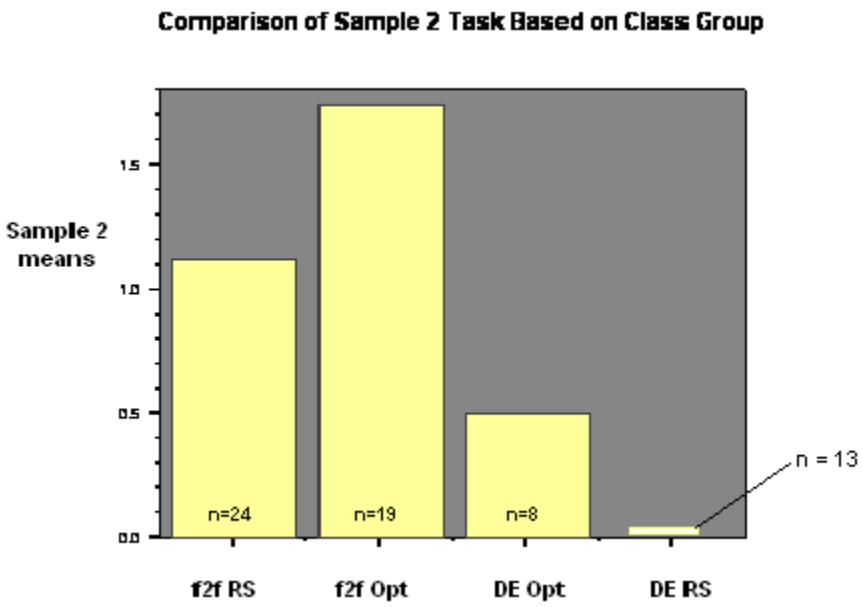


Figure 10: Comparison of Two Mean Scores

The scale on the above table begins at zero and is much smaller than the sample one scale. This is due to the fact that many students were unable to perform this speaking

task. In fact, none the students in the DE required speaking section was able to produce any measurable speech (they reported simply not being able to answer the prompt).

They sent WAV files with statements similar to "I cannot respond to this right now." Similar to the Sample Item One mean scores, Sample Item Two mean scores established that all four groups began with similar speaking abilities.

With the validation that there were no significant speaking differences among the four groups at the onset of the course, it creates the possibility to look to delivery modes and speaking amounts during the semester to explore changes, if any, that might result by the end of the semester.

RESEARCH QUESTION ONE: EXPLORING PERFORMANCE SCORES ON SPEAKING ASSESSMENTS (REPEATED PROMPTS)

This question addresses the comparison of overall performance on four speaking prompts that were presented on the final speaking assessment (time 2 measure). All four of the speaking prompts were either first introduced at orientation or at the midterm (time 1 measure). The specific research question is: Are there differences among the sections' performances, regardless of delivery mode and assigned speaking tasks, on the same speaking tasks recorded at two points in a semester?

Orientation Item One to Final Item One Speaking Performances Compared

The sample speaking prompt 1 scores (Time 1) were paired with final speaking prompt 1(Time 2). Table 9 presents the descriptive statistics: mean scores (*M*), standard deviation (*SD*) and number in each subgroup (*n*).

Table 9
Descriptive Statistics for Sample Item One to Final Item One Performance Scores

		f2f Required	f2f Optional	DE Optional	DE Required	Total
Sample 1	**M**	8.08	8.95	7.88	8.54	8.41
	SD	5.241	5.788	5.842	4.054	5.166
	N	24	19	8	13	64
Final 1	**M**	12	11.53	11.63	13.77	12.17
	SD	2.396	2.412	3.15	1.73	2.479
	N	24	19	8	13	64

There was no statistically significant main effect for class [F(3, 9.045) =.4.95, p>.05]. The effect size was medium (partial eta squared = .24). For this speaking task, there were no significant differences among the four class mean scores. There was a significant effect of time [F(1,60) = 26.420, p < 0.05]. There was no significant interaction. The class does not moderate the time effect [F(3,60)= 0.601, p = 0.617]. Each class mean score improved from time 1 to time 2 but the increase is not due to class features (DE or Required/Optional Speaking).

Most students were able to attempt the sample one prompt. The f2f optional group scored the highest at the beginning of the course, but this high score gave way to the DE group with required speaking (see Figure 11).

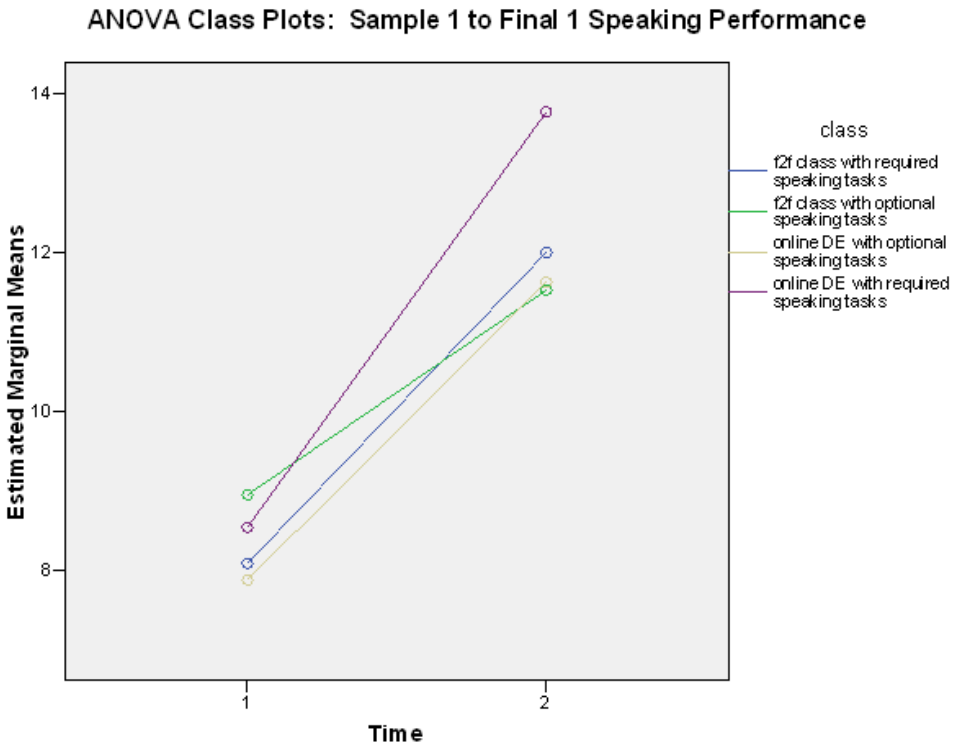


Figure 11: ANOVA Class Plots: Sample Item One to Final Item One Speaking Mean Scores

This figure shows four mean scores above time 1 that correspond to the mean scores from Table 9. The lines represent the gains in mean scores from time 1 to time 2. The high scores from the final performance had the two required speaking classes above the optional speaking classes. The DE required speaking class had the top mean score, this suggests, for this speaking item, mode of course delivery had very little to do with achieving the high score.

Orientation Item Two to Final Item Two Speaking Performances Compared

A two-way ANOVA between group analysis of variance was conducted to explore the impact of class on the speaking performance scores between the initial second sample and the second final speaking performances. There was no statistically

significant effect for class [$F(3,60)=1.169$, $p=0.329$]. Table 10 presents the descriptive statistics.

Table 10
Descriptive Statistics for Sample Two to Final Two Performance Scores

		f2f Required	f2f Optional	DE Optional	DE Required	Total
Sample 2	**M**	1.13	1.74	.50	0.00	1.0
	SD	3.248	3.263	1.414	0.00	2.73
	N	24	19	8	13	64
Final 2	**M**	10.50	10.00	7.63	10.92	10.80
	SD	3.284	2.728	4.838	3.068	3.382
	N	24	19	8	13	64

The classes each made improvement from time one to time two [$F(1,60)=343.427$, $p<0.05$]. The DE Required Speaking group could do nothing but increase in their performance with an initial mean score of zero. At the end of the semester the top score for three classes were very close. However, the DE Optional Speaking class' mean score is much lower than the three higher clustered groups (see Figure 12). Although the final differences were not significant, this is the second speaking performance when the DE Optional class has the lowest mean score.

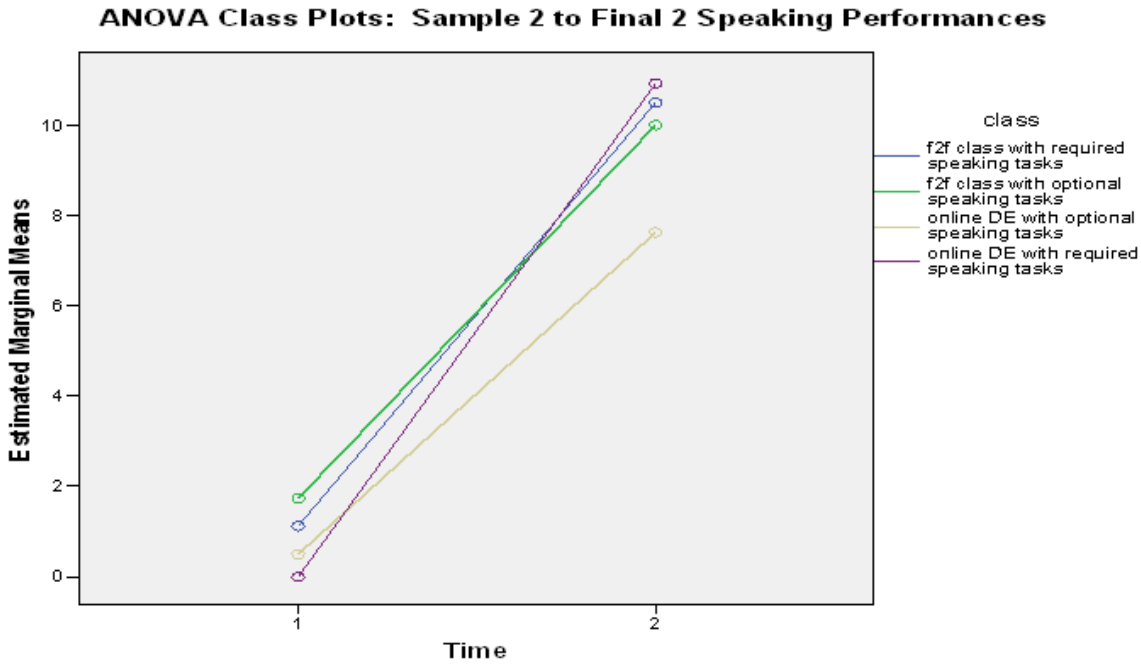


Figure 12: ANOVA Class Plots of Sample 2 to Final 2 Speaking Performance Mean Scores

The above figure shows the four classes mean scores from the orientation sample two located at Time 1. The mean scores correspond to the scores from the descriptive statistics Table 10. The lines indicate the increase in mean scores to Time 2 taken at the final speaking assessment. All sections made improvement but the differences of the final mean scores were not significant in a between class comparison.

Midterm Item One to Final Item Three Speaking Performances Compared

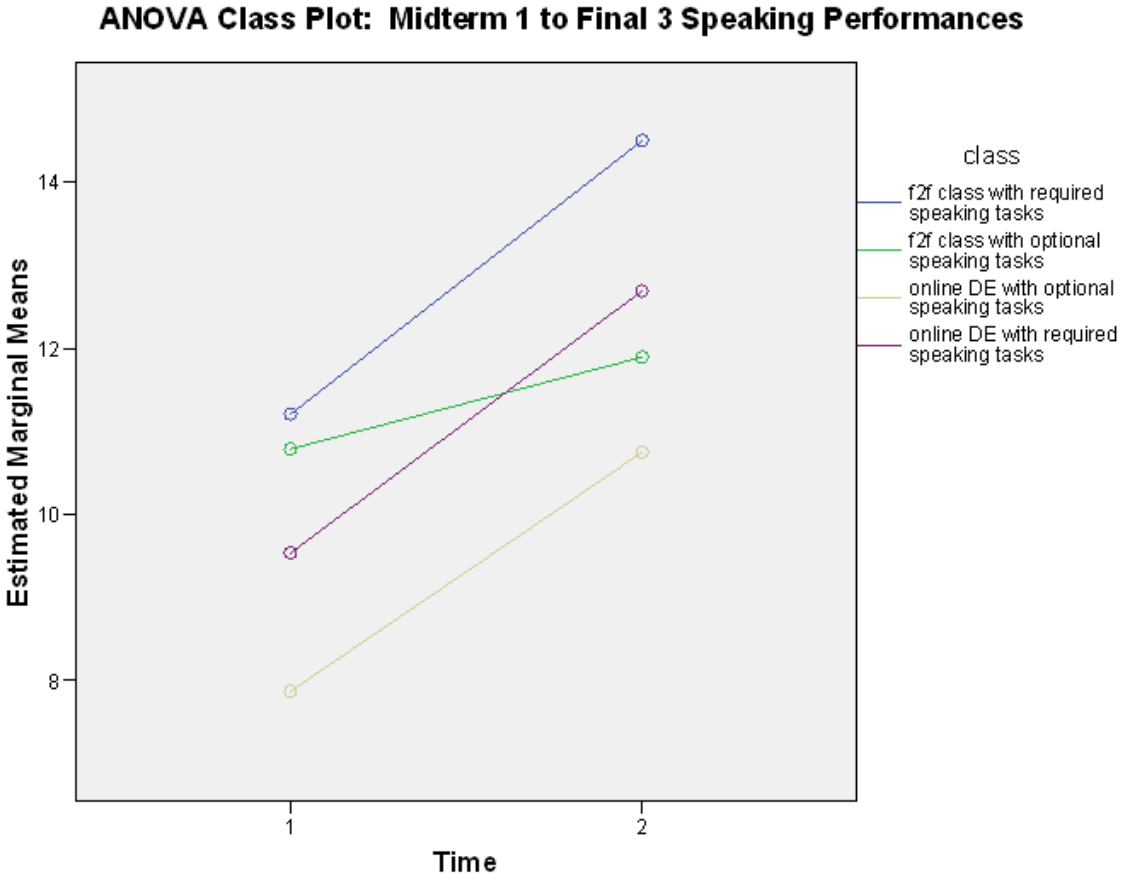
A two-way repeated measure ANOVA was conducted to explore the impact of class design on speaking performances between midterm speaking prompt 1 (Time 1) and final speaking prompt 3 (Time 2). The means and standard deviations are presented in Table 11.

Table 11
Descriptive Statistics for Midterm Item One to Final Item Three Performance Scores

		f2f Required	f2f Optional	DE Optional	DE Required	Total
Midterm 1	**M**	11.21	10.79	7.88	9.54	10.33
	SD	3.659	4.049	4.824	3.929	4.044
	N	24	19	8	13	64
Final 3	**M**	14.50	11.89	10.75	12.69	12.89
	SD	1.474	2.846	4.062	3.172	2.955
	N	24	19	8	13	64

Both within and between measures analyses show improvement. There was a significant effect of time [$F(1,60)=35.55$, $p < 0.05$]. This means that there were improvement for all class sections between the midterm mean score and the final mean score for the same speaking prompt. There was also a main effect of class [$F(3,60)=3.12$, $p = 0.032$]. This means that the score between the sections were significant. Post hoc comparisons using the Tukey HSD indicated that the difference man score between f2f Required Speaking and the DE Optional Speaking was $M=3.54$ at a significance level of $p<0.05$. There were no other significant mean differences reported. Figure 13 shows the changes in oral performance mean scores between the midterm (Time 1) and the final (Time 2).

Figure 13
Title: ANOVA Class Plot: Midterm 1 to Final 3 Speaking Performances Mean Scores



The figure above shows the distribution of the midterm mean scores (Time 1) by class with the highest mean score for the f2f Required speaking followed by the f2f Optional speaking, in third position DE Required speaking and the lowest mean score DE Optional speaking. Each class line slopes in a positive direction indicating an increase in the mean score between the midterm and the final. This increase was significant for each section.

The difference in the final mean score between the f2f Required speaking class and the DE Optional speaking class is clearly indicated above Time 2. This difference is significant. The results from this comparison suggest that class delivery is not as

important as required speaking practice. A detailed qualitative analysis of the specific differences will be addressed in research question two.

Midterm Item Two to Final Item Five Speaking Performances Compared

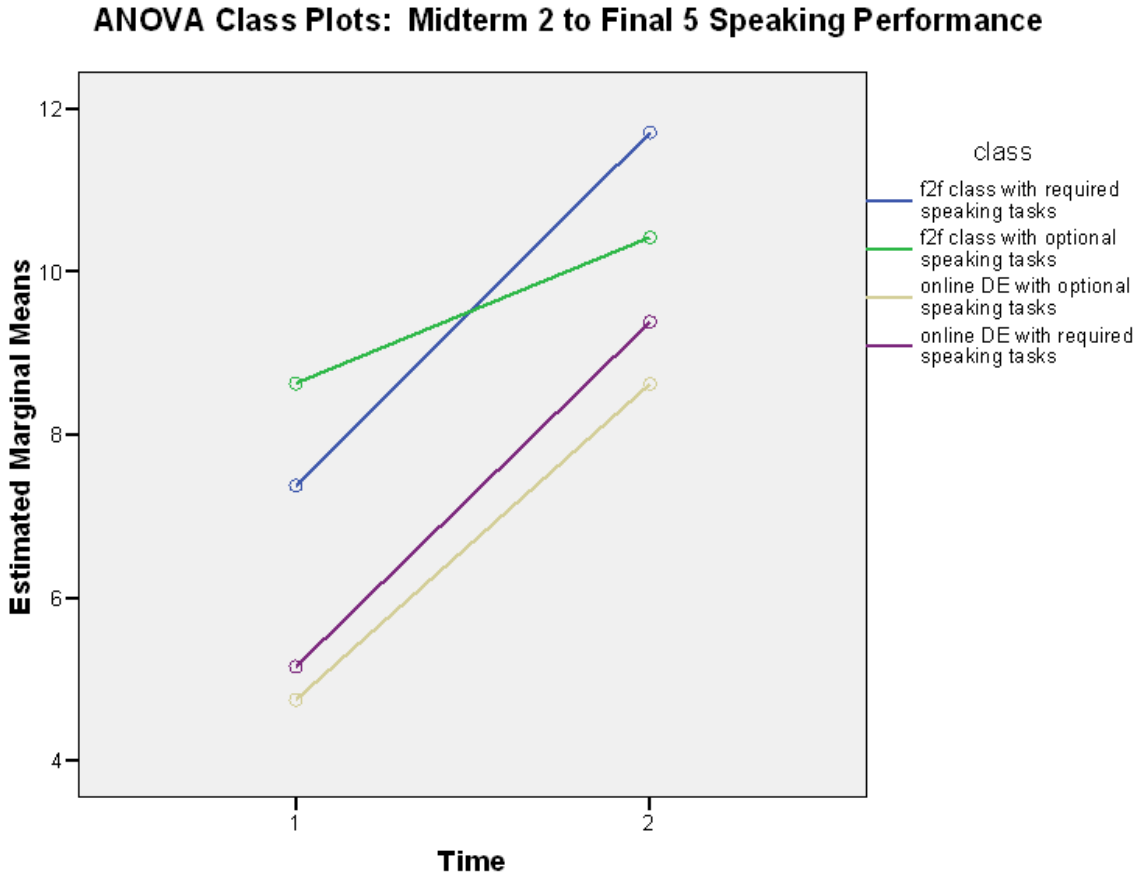
The fourth ANOVA statistical procedure was used to analyze speaking performances between midterm speaking prompt 2 (Time 1) and final speaking prompt 5 (Time 2). The means and standard deviations are presented in Table 12. Although there was no significant effect based on class design [$F(3,60) = 2.76$, $p = .089$], there was significant difference based on time [$F(1,60)=3.02$, $p =.037$].

Table 12
Descriptive Statistics for Midterm Item Two to Final Item Five Performance Scores

		f2f Required	f2f Optional	DE Optional	DE Required	Total
Midterm 2	**M**	7.38	8.63	4.75	5.15	6.97
	SD	4.362	4.705	4.683	4.562	4.677
	N	24	19	8	13	64
Final 5	**M**	11.71	10.42	8.63	9.38	10.47
	SD	2.836	2.815	3.739	3.709	3.261
	N	24	19	8	13	64

Each class improved from the midterm to the final. There were no significant differences found among class mean scores. Figure 14 shows the mean scores for all four classes for the midterm (Time 1) and the final (Time 2).

Figure 14



The above figure demonstrates the relative increase from the midterm mean scores on the left of the slope to the final mean scores on the right of the slope. In this case, both f2f classes score higher than the DE classes. But it is interesting to note that the f2f class with the initial higher score (optional speakers) did not maintain a top score on the final. The f2f required speaking class outperformed the f2f optional class while the DE required speaking class maintained its higher position over the DE optional speaking class.

Chi-Square Report

Overall, the performance scores of the participants did not differ by class, but in the case of one speaking task – the third speaking task on the final exam. The following are the reported chi-square statistics with the alpha level of significance (0.05) (see Table 13).

Chi-Square Reports for the Final Speaking Items	
F1	$x^2(30,N=66) = 31.79, p = .38$
F2	$x^2(39,N=66) = 39.88, p = .43$
F3	$x^2(33,N=66) = 45.19, p = .047$
F4	$x^2(33,N=66) = 43.03, p = .113$
F5	$x^2(39,N=66) = 44.25, p = .269$
F6	$x^2(39,N=66) = 44.36, p = .256$

Table 13: Chi-Square Reports for Final Speaking Items

This test compares observed data with what we would expect to get (if the null hypothesis of no difference was true). It is based on the principle that if the two variables are not related (for example class is not related to score) then measures applied to each variable will give similar results (for example about the same proportion of each class would have similar scores), with any variation between the results being purely caused by chance. If the experimental measures are significantly different, then some relationship can be claimed. This is the case only for one final speaking item: Speaking item three which will be discussed further in research question two. Speaking items four and six will be discussed further in research question three.

Discussion Research Question One

Are there differences among the sections' performances, regardless of delivery mode and assigned speaking tasks, on the same speaking tasks recorded at two points in a semester?

Findings & Discussion

The first finding is that for each of the four repeated measures, each class made an improvement between Time 1 and Time 2 performances. This had not been documented before. Perhaps this is because, as Blake (2008) suggests, that proficiency measures are not sensitive enough to capture small developmental gains that beginning L2 learners exhibit. This study, however, did not use generalized proficiency measures of the OPI, but rather expanded the scales of Borrás and Lafayette (1994) and Koren (1995). The mean scores based on this rating scale allowed for a more sensitive outcome.

Another reason for the significant improvement might be due to the materials. Perhaps they provided comprehensible input (Krashen, 1981). Students were able to interact with the materials via CD-ROM, web pages and for the DE students, an interactive eBook.

A third argument for improvement for all classes might be attributed to the repetitiveness of the assessment and the exposure to similar speaking tasks during the semester. Even students not performing the speaking files in the optional speaking sections still had access to view the prompts.

A fourth argument for all class improvement is that only course completers performed on the final speaking assessment. Those students who did not keep up with the course or were performing poorly withdrew. Student withdraws were high for the DE sections after the midterm exam (week 9) and after Thanksgiving (week 12). The withdrawal rate was about average for the f2f Optional section. The retention rate was high for the f2f Required Speaking class.

The second finding is mixed. One of the three *between group* analyses yielded a significant difference between only two of the classes: f2f Required Speaking and DE Optional Speaking. With five of the six speaking performances among the groups displaying no significant difference, this suggests that the distance learning delivery mode can produce equivalent outcomes and as such, is a viable learning environment for foreign languages. This study adds to the few speaking studies finding there is no difference with DE and f2f L2 learning (Cahill and Cantanzaro, 1997; Despain, 2003; Blake and Delforge, 2005; Blake, 2008) but with the caveat that required speaking tied closely to curricular objectives be included.

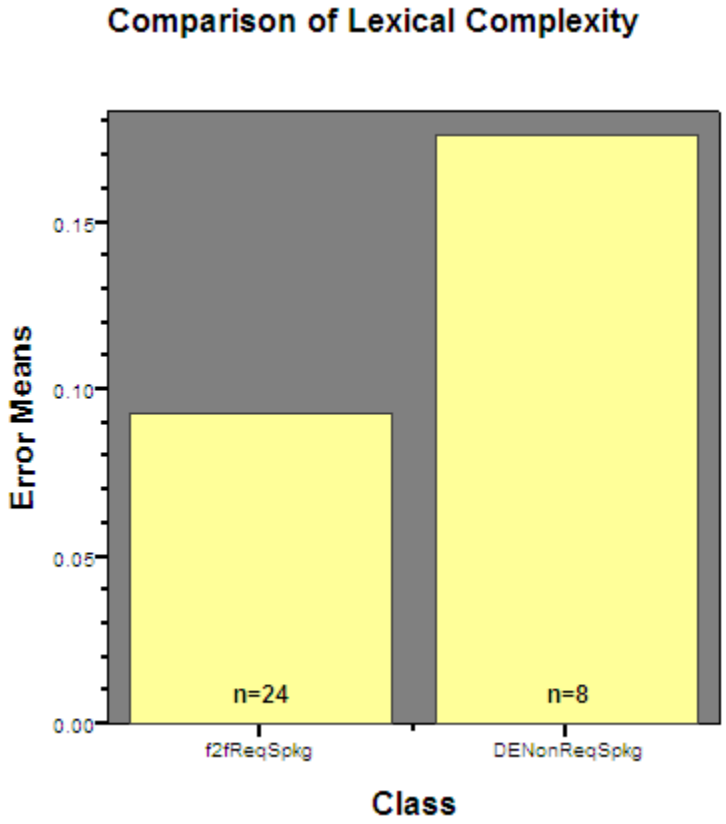
RESEARCH QUESTION TWO: DIFFERENCES

If there is a difference between delivery mode oral performances, what is the nature of the difference?

The second research question concerns the nature of the significant difference between midterm one and final three speaking assessment scores for the f2f Required class and the DE Optional class. Having established this difference, the WAV files were transcribed and evaluated for lexical density, lexical complexity and Target Language Use (TLU) for inflectional words.

There were no differences found for lexical density between the two groups ($F(1,31)=1.32$, $p>.05$). Significance was found between groups for lexical complexity ($F(1,31)= 4.28$, $p<.05$). Figure 15 shows the comparison of lexical complexity means.

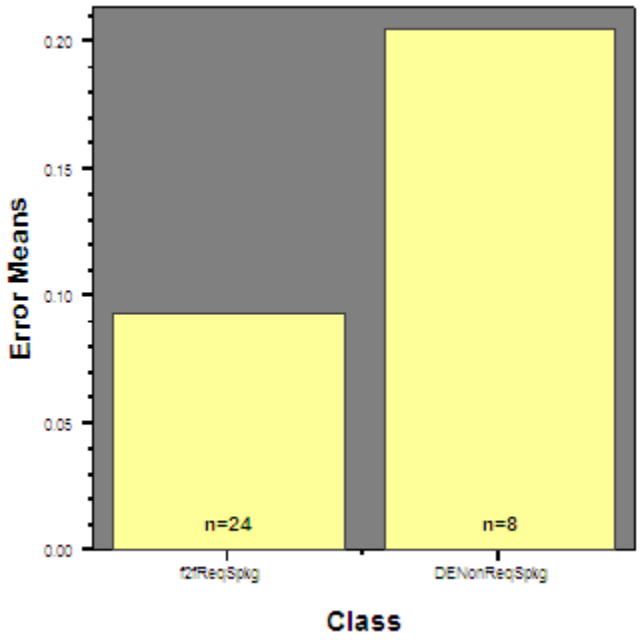
Figure 15



The above bar graph shows the f2f required speaking group with more variety and richness in the vocabulary spoken on final speaking item three. The DE optional group had a smaller variation or range of vocabulary. The formula gives us an inverse relationship: the broader the range of vocabulary, the lower the score and vice versa; therefore, a lower score indicates a richer lexicon in the speaker (see Appendix G).

The following bar graph displays the means for the two classes. The higher mean shows more errors and was significantly different. TLU errors for inflection was also significant ($F(1, 31)=4.34$, $p<.05$) with a large effect size (partial eta squared = .12) The group with distance education with optional speaking produced more errors in their output. Figure 16 displays the comparison of TLU errors.

Figure 16: Comparison of TLU Mean Errors



The above bar graph (Figure 16) shows the error of inflectional word use. The DE Optional speaking class made significantly more errors than the f2f Required speaking class.

Findings and Discussion

The significant difference discovered in the repeated measures and further analyzed discovered that there were more grammatical and lexical errors made by the DE Optional Speaking class.

The above results about accuracy seem to contradict the findings for CMC (e-mail) research. Using CMC in e-mail exchanges, some researchers found an increase or difference in lexical density rather than accuracy (Gonzalez-Bueno, 1998; Aitsiselmi, 1999). Other researchers suggest that by participating on a regular basis in networked-

based classroom activities simulating "speaking" tasks, there is a promise for gains in TL proficiency development, lexical development or the production of idea units (Blake, 2000; Payne & Whitney, 2002; Abrams, 2003).

Another aspect of task performance may be intertwined with overall task difficulty (Skehan, 1998; Iwashita, McNamara & Elder, 2001) or on particular aspects of performance (Rahimpour, 1997; Robinson, 1995). Wigglesworth (1997) investigated the impact on planning time of measures of accuracy, fluency and complexity with mixed results. The current results also contribute to mixed findings.

RESEARCH QUESTION THREE: NEW SPEAKING PROMPTS ON ASSESSMENT

How do the different class sections compare on the final speaking exam when new testing prompts are introduced?

The first research question dealt with repeated measures. The final speaking assessment also included two new speaking prompts. This question investigated how each class performed relative to each other on the two new prompts.

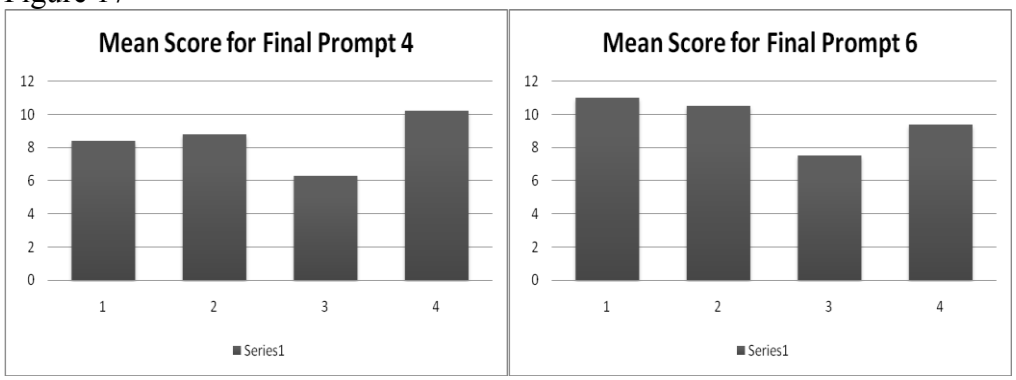
A one-way analysis of variance (ANOVA) between groups was used to compare the groups' performances on Final Item Four and Final Item Six speaking prompts. Neither speaking score analysis showed statistically significant differences between classes at the $p<.05$ level; The Final Item Four speaking performance was [$F(3, 60) = 1.7$, $p = .175$] and the Final Item Six speaking performance was [$F(3, 60) = 1.5$, $p = .225$]. The effect size calculated using the eta squared was .078 for the Final Item Four and .070 for the Final Item Six prompt. Post-hoc comparisons using the Tukey HSD test indicated that mean score was not statistically significantly different on either prompt.

The following, Table 14 and Figure 17, demonstrate the score means for each class. On the two final speaking prompts, the DE optional speaking class' mean score is the lowest.

Table 14
Descriptive Statistics for Final Item Four and Final Item Six Performance Scores

		f2f Required	f2f Optional	DE Optional	DE Required	Total
Final Four	**M**	8.46	8.79	6.25	10.80	8.61
	SD	3.647	4.117	3.284	3.840	3.857
	N	24	19	8	13	64
Final Six	**M**	11.00	10.47	7.50	9.38	10.08
	SD	4.294	3.289	4.342	5.455	4.347
	N	24	19	8	13	64

Figure 17



Legend: 1 = f2f Required 2 = f2f Optional 3 = DE Optional 4 = DE Required
This figure to the left shows the mean oral performance scores for prompt four.

Although there was no significant difference, the DE Optional Speaking class continues

to be placed with the lowest score. The DE Required Speaking class had the highest mean score.

This second figure, on the right, shows the mean oral performance scores for prompt six. The DE Optional Speaking class continues to underperform the other class sections.

Findings and Discussion

The data shows no significant difference between speaking levels and course delivery modes. It further shows no difference between classes when new speaking prompts are introduced as opposed to repeated speaking prompts.

Side to side, the mean scores demonstrates variability in oral performance for different speaking prompts which may be due to task variables. These findings support the research on task variation, primarily task difficulty (Skehan, 1998; Iwashita, McNamara & Elder, 2001). When the task has more competing components, i.e. simultaneous steps to complete, performances levels vary.

RESEARCH QUESTION FOUR: SELF-ASSESSMENTS AND STUDENT BELIEFS

What do students report on their final exam self-assessment in terms of their own performance confidence levels and beliefs about learning activities? Are there any differences in opinions among the delivery modes and oral activity amounts?

Research question four investigated post-final assessment self-report data. There were two performance self-rating data: a numerical assignment to "confidence" level for each speaking prompt and open-ended comments regarding the speaking prompt.

Confidence Categories

Following the submission of the final speaking assessment, the exam proctor(s) provided each student with a self-assessment and reflection form (see Appendix E). Students were asked to self-rate their own performance based on how well they thought they performed with low (1, 2) to high (3, 4). Admittedly, this is a difficult construct to pin down. The following Table summarizes by class section how the majority of students believed they performed.

Table 15:
Confidence Score of Performance for Each Final Speaking Prompt

% of High Confidence Level	f2f Required Speaking n=24	f2f Optional Speaking n=19	DE Optional Speaking n=8	*DE Required Speaking n=13*
Final One High	66.7%	68.4%	62.5%	*61.5%*
Final Two High	62.5%	68.4%	75%	*53.8%*
Final Three High	66.7%	84.2%	62.5%	*46.2% **
Final Four High	20.8%	10.5%	28.5%	*7.7%*
Final Five High	58.5%	68.4%	62.5%	*38.5%**
Final Six High	*25%*	*42.1%*	*50%*	*38.5%*

* Contrary to other classes' high % self-scores

The table shows the percent of students within the class that gave a self-rating score of 3 or 4. The final speaking prompts are labeled by item number from prompt one to prompt six on the last row. The class percents run in columns.

There are two observations that can be made from the table. The first observation is that for *new* speaking prompts, all class sections had fewer students with

high confidence for new test prompt four. Speaking prompt six also contained fewer students reporting confidence on their performance.

The second observation is that the DE Required Speaking student appear to be the least confident with fewer student self reporting high confidence. The DE Required speaking group of students had fewer students self-reporting high on speaking prompts three and five; both of which were repeated prompts from the midterm. This is in contrast to the other three groups of students with most students reporting high confidence on their performances.

Self-Report Comments

Half of the students from the f2f groups (21) did not provide any comment after the self-rating on the range 1 – 4. Only three students from the DE groups provided comments. Those that did make comments chose to do so to cite frustrations with a) cognitive overload, b) grammatical/lexical shortcomings, and 3) anxiety. Several students described the pre-final preparation (a study strategy) and reported success.

Discussion & Summary

There is already established evidence that "DL students are recalcitrant to responding to requests for cooperation outside the boundaries of class assignments" (Blake, 2008). There seems to be evidence that this lack of response is also associated for requests not directly responsible for grades as well. Because there were only three DE comment responses, the following reflects the entire four sections of test takers' comments. The comments from the three DE students are similar in content to the f2f student comments.

In regard to cognitive overload, several comments cited not remembering specific words, phrases or conjugation forms. "All I could remember was how to say *siento* (I think that is I feel)." This is a good example of partial retrieval. Another series of comments directly addressed vocabulary retrieval. "Suddenly my vocabulary seemed to shrink down to nearly nothing." "I completely went blank on the days of the week." "I just couldn't remember." And an emoticon was used, "☹ Mind went blank." These comments relate directly to Levelt's (1989) model of speech processing: especially formulation (drawing on the lexicon and later the grammatical features) and internally formulated utterance and articulate. Because Levelt dealt with "time" pressure, it appears that students knew what they could not retrieve within their test taking time on task. In Bachman's model (1990), the students' ability breaks down. The student cannot use his grammatical competence to retrieve vocabulary.

Another set of breakdown surrounded TL usage. Grammatical accuracy was noticed. One student reported that "I think I remembered the correct verbs, but not sure of the conjugation. Another student discloses her lack of short-term memory and inability to store lexical items. "I am very good at conjugation and sentence construction. However, my vocabulary is terrible. I have great difficulty with that. I had the same problem back in high school." There seems to be an awareness of memory issues but this student lacks of strategies to overcome those shortcomings.

It was interesting to note two comments concerning interpretation of instructions. "I had a hard time with the pronunciation of how I would feel **if** something happened." This prompt asked them to describe how they feel when a particular situation occurs. One student noted that she was unsure about the ambiguity of the prompt. "I was

unsure if the schedule included anyone other than the student so I used *vas* for each sentence." This student demonstrated pragmatic competence with her uncertainty to interpret the speaking prompt.

There were frustrations with taking the task too literally. "I really don't do much on the weekends. So this was a difficult one to speak about." The student did not feel a connection with the prompt and found it difficult to create a response with an inauthentic response or could not reach into activity vocabulary to use daily activities.

One f2f student decided to evaluate the tasks as "hard, fairly simple and eeya…lol." This was interesting because it was a speaking test and the "eeya" was a transcription of a vocalization to represent difficulty. Another student commented that, "All of the tasks will be used in real life situations and were helpful." This *perhaps* implies motivation for performance. It did validate the intention of the speaking prompts as a test design feature.

Time was an issue. Two students commented that they ran out of time. One DE optional speaking student became frozen. She comments, "…as you can see, I didn't submit anything for it." One set of responses were self critical. "I skimped on my answers." "I don't know if I was able to answer the question completely."

One f2f student commented discussed final exam preparation. "I studied the last chapters & it was like the first few clipped out of my brain in the speaking tasks." Because she was focused on the later part of the course content, she did not focus on the content presented and practiced toward the end of the semester: content introduced on the repeated speaking items. Another comment, from the DE required speaking group, had to do with comfort at having remembered the exact task: "I remember doing that last

time so it helped." It is interesting to note that none of the repeated prompts had repeated topics from students. For example, no one used the same family member they had described on the midterm prompt on the final prompt. This was true across class sections.

Not all comments were based on frustration. Regarding repeated speaking prompts, one student remarked that "I really improved on this from the first of the year." Other comments were less enthusiastic, "I think I did okay." Here are two very proud comments: "I finally know what I'm saying." "I think I completed this speaking file nicely." These comments further support the dimension of task repetition within the context of language testing (Lee, 2000). In addition, task repetition may have the ability to promote a sense of confidence and comfort level.

The comments provided by student post-final speaking exam reflections provide insight to the students' view of the testing experience. It revealed memory concerns, self-reported grammatical/lexical shortcomings, anxieties, pre-exam preparation and successes. The next set of student comments reveal what semester activities are believed to have been useful in promoting speaking.

Students Selection of Helpful Materials: (Research Question Four Continued)

There was an open-ended comment following the self-rating and comments concerning confidence. The question asked students to identify activities that were most helpful in contributing to the student's speaking skills during the semester.

Comments students provided were analyzed and categorized by class section. Table 16 shows the frequency count and percent of total times the activity was mentioned as useful promoting or assisting in Spanish speaking skills. Because this was an open-

ended question, a student may have mentioned more than one activity. Therefore, the percentages are in ratio to all activities.

Table 16
Student Identified Helpful Activities for Speaking

Activity	f2f Required Speaking 24 students	f2f Optional Speaking 19 students	DE Optional Speaking 8 students	DE Required Speaking 13 students	Activity Totals
WAV files	13	1*	0	3	17 **23.9%**
In-Class /On-line Speaking	13	10	4	3	30 **42.2%**
Listening Lab	3	2	2	1	8 **11.3%**
Writing Workbook	2	0	1	1	4 **5.6%**
CD-ROM	1	1	2	2	6 **8.5%**
Online Websites	0	1	0	0	1 **1.4%**
eBook	NA	NA - 1**	1	1	3 **4.2%**
Tutoring	2	0	0	0	2 **2.8%**
Total by Class	34 **(47.9%)** 100%	16 **(22.5%)** 100%	10 **(14%)** 100%	11 **(15.5%)** 100%	71 **100%**

*** Available for feedback- 1 student comment regretting not having done them. **Not available to individual students. Available in class only: 1 student comment wishing for personal access.**

Based on these responses, it was interesting to see how much more "participation" overall was activated by the f2f required speaking class. They not only had the highest retention rate, but provided nearly half of the comments. This may suggest that online

"engagement" measures or use of course study tools and materials (assigned/unassigned) might be worthy of future research (White, 1995).

Of the total 71 activities identified, the students in the f2f Required Speaking class, provided 47.9% of the total activity identifications. This was followed by the f2f, Optional Speaking class at 22.5%, the DE speaking at 15.5% and DE Optional Speaking class at 14%. The f2f Required Speaking class selected the WAV files as the most helpful activity to improve their speaking Spanish; however, this only represented 23.9% of the total comments. This is also a natural because this type of activity was required and students were motivated for grades to comply with the assignment. Students wrote that "WAV files were less stressful than being in from of the class and speaking." This provides additional evidence of lowering stress while requiring all to participate.

Selected as the most helpful in improving speaking during the semester was *in class and online class speaking* at 42.2% of the 71 comments made. The f2f required speaking class tallied 13 references or 43.3% of their group's total. F2f optional speaking class identified synchronous speaking in class 10 times or 33.3% of their group's total followed by DE optional speaking with four comments or 13.3% and DE required speaking with three comments or 10% of their group's total. All students were encouraged to speak with each other online in a synchronous way. Even though this was not monitored in the classroom closely or online at all except in a teacher-lead class meeting.

The third most identified activity for helping speaking skills were the listening activities at 11.3% of the total 71 comments made. The listening lab activities done together in class with the f2f required speaking class was identified at 37.5% with three

comments. This is followed by 25% (2) for the f2f optional speaking class, DE optional speaking class at 25% (2) and DE required speaking class at 12.5% (1).

The most surprising were the few mentions of the CD-ROM, written workbook, eBook, tutoring, and online websites activities as helping improve speaking skills at 9% each at 8.5%(6), 5.6%(4), 4.2% (3), 2.8%(2) and 1.4%(1) respectively. Except for the written activities from the workbook and the required CD-ROM reports for all classes, all other activities were unmonitored. This suggests that required and graded activities may be perceived as more valuable in language learning than other "non-essential" activities.

One reason for identifying the CD-ROM as more helpful in developing speaking skills is that it contained a voice recorder and enabled students to compare their voice with a native speaker voice. The recording could be converted to a WAV file and saved for later listening. This was not identified as very important or perhaps this type of CD-ROM activity was not performed regularly.

SUMMARY

The student profiles and initial performances show no main differences among the four beginning Spanish classes at the onset of the course. The majority of students arrived to the course with previous formal experiences in learning Spanish. Moreover, the majority selected speaking in Spanish as the primary goal.

Although the speaking performances of three of the four paired analyses showed no differences among the four classes, the finding of no significant difference between classes that have been exposed to different class treatments has been found in previous studies (Lumley & O'Sullivan, 2005; Shimazu, 2005; Blake, Wilson, Cetto & Pardo-

Ballester, 2008; Villas Rogers, 2002). This study supports Wigglesworth (1997) who found mix results with planned time in language-testing contexts.

When using the same one-way ANOVA procedures and applying it to compare all six final speaking performances, the only statistically significant difference was found on Final Item Three speaking mean scores between the DE optional speaking class and the f2f required speaking class. The specific differences were found in lexical and target language usage scores. The results about accuracy seem to contradict the findings for CMC some researchers found an increase or difference in lexical density rather than accuracy (Gonzalez-Bueno, 1998; Aitsiselmi, 1999).

Yuan and Ellis (2003) suggest that relationship between planning and L2 oral output are interrelated and are of pedagogic relevance as they indicate the task conditions needed to promote accuracy, complexity and fluency in a monologic speech production. With one performance difference between course delivery mode mean scores, the findings suggest that planning is an important element for beginning speakers. Similar findings about advanced planning for complexity are suggested by Ortega (1989). Crookes (1989), Foster and Skehan (1996), Wendel (1997), Mehnert (1998), and Ortega (1999) all report that pre-task planning results in greater complexity of language.

The self-report data introduce language learner variables: anxiety, learner strategies, motivation and engagement. These comments and topics further support the dimension of task within the context of language testing (Lee, 2000) and online distance education activity design (Hurd & Ortega, 2001; Ariza & Hancock, 2003; Levy, 2003; Doughty & Long, 2003; White, 2005). Based on student self-reports and reflection,

future research should include online language student engagement and overall online language course anxiety and online foreign language speaking anxiety.

A summary of the results, a discussion of the implications these finding have for practical implementation of FL DE learning, the limitations of the current study, suggestions for future research, and offer some concluding remarks will be presented in Chapter Five.

Chapter 5: Conclusion

INTRODUCTION

This study was designed to compare courses within two learning environments: the physical classroom supported by language tools and materials using today's technologies, and the virtual classroom, with today's course management systems and online language materials and tools. Specifically, it sought to investigate speaking outcomes with two levels of assigned speaking tasks. It explored students' views of the speaking assessment and materials. This chapter will summarize significant findings, provide theoretical and pedagogical implications, make general recommendations, discuss the study's limitations, discuss future research and offer some concluding remarks.

OVERVIEW OF SIGNIFICANT FINDINGS

Research Question One

> ***Are there differences among the sections' performances, regardless of delivery mode and assigned speaking tasks, on the same speaking tasks recorded at two points in a semester?***

There was significance reached in one of four final speaking prompts. The difference was between the f2f Required Speaking group and the DE Optional Speaking group. For the other three final speaking prompts, there was no significant difference of student performances between course delivery mode and required speaking task amounts. There was a large amount of input provided by the course materials which may suggest that Krashen's comprehensible input hypothesis is a viable component to language output.

Research Question Two

If there is a difference between students' performance as a result of the different delivery modes, what is the nature of the difference?

The significant difference between the f2f Required Speaking and the DE Optional Speaking groups were obtained on the final speaking assessment prompt three. The specific difference was in lexical and morphological error counts. The DE Optional Speaking section made significantly more errors in lexical complexity and TLU counts.

There are two plausible explanations for the differences between the f2f required speaking group and the DE optional speaking group. The f2f students brought many family pictures to class during that chapter and were very excited about talking about family. There were three siblings in this one section that created "fake" stories about each other. Also in this class, there was a mother-daughter study team. During class, these family members were separated from each other to practice with other students. Students in this class had speaking practice in class and at home in addition to the chapter speaking task sets.

The DE optional students provided family descriptions in writing in the discussion board. No one in the DE optional speaking group chose to attach a picture. The DE required speaking group had several students with pictures of themselves and family members. (Notes taken from the instructor's journal.)

The speaking task itself did not require any more difficult or less difficult language processes for interpreting the task instructions. Because it was a repeated speaking task from the midterm assessment, it is difficult to isolate an associative variable to support the difference.

Research Question Three

How do the different class sections compare on the final speaking exam when new testing prompts are introduced?

There was no significant difference among the four classes. There were lower mean scores for the DE Optional Speaking group on both new test prompts. Final speaking item four provided the classes with the most challenge. Comments provided after the exam revealed that the image, used to support the schedule, contained too many ideas. The final speaking prompt four mean score was the lowest for each class.

Research Question Four

What do students report on their final exam self-assessment in terms of their own performance confidence levels and beliefs about learning activities? Are there any differences in opinions among the delivery modes and oral activity amounts?

The DE Required Speaking section consistently self-assessed with a low confidence level. The other three sections were confident with their final speaking performances. This did not include when new speaking prompts were introduced. When the new speaking prompts were introduced, all sections fell in their self-rating confidence level. There were only three of 21 comments provided by DE students in terms of their performance on the final speaking test. For those that did respond with comments regarding oral performances, they identified (1) cognitive overload, (2) grammatical/lexical shortcomings, and (3) anxiety.

There was a high participation rate for the f2f Required Speaking group for identifying activities they believed helped develop speaking skills throughout the semester. These were the activities as most having help developed speaking skills: (1)

in-class and online speaking activities, (2) creating WAV files, and (2) listening activities. It was interesting to note that not all students perceived learning tasks equally.

THEORETICAL & PEDAGOGICAL IMPLICATIONS

Speaking (communicating with voice) is one of the most challenging skills to evaluate in the regular routine of conducting class and providing feedback. The following theoretical implications can be drawn from the series of analyses completed with the speaking assessment files in regard to approaches to activity design. Research has shown that focus on meaning via pushed output is not sufficient for students' to reach native-like productive skills (Swain, 2003). Comprehensible input is necessary, but not sufficient (Doughty & Long, 2003).

The best results are not with kill & drill discrete point grammar teaching where students spend most of their time working on isolated linguistic structures (However, some discrete point work is not out of the question either). This study suggests that those students who participated in meaning-focused speaking activities as part of the regular "class" and required additional meaning-focused speaking activities outperformed in number but not signficantly those in the DE section who chose not to do those activities as assigned (but unmonitored) or optional. The statistical scores revealed that the difference was not significant. In the early skill building phases of speaking, focus on form activity tasks help pull the discrete points across the chapter by focusing on meaningful communication.

Speaking Activities and Assessments for Beginning Spanish

The additional speaking WAV files were necessary for the online DE courses as the only way to provide offline speaking. DE students who participated in the speaking

section not only had a slightly higher retention rate, but scored higher on the DE section with optional speaking. Requiring speaking for DE is essential to motivate students to participate and perhaps motivate students to remain engaged in the course.

I hypothesized that students with WAV file practice would feel more confident with their performances. This turned out not to be the case as the DE Optional speaking section was confident with the final speaking files created while the DE Required speaking section was less confident on all WAV files presented at the final. The research with online FL speaking anxiety needs further exploration.

GENERAL RECOMMENDATIONS

Moving Narratives to Interactive Synchronous Speaking

I believe that the by introducing offline speaking tasks that bridge several discrete points (vocabulary/grammar) and are communicatively situated, created opportunities for practice students in DE online first year FL/SL courses that were not previously required in earlier semesters. Providing feedback to required work is time consuming but motivating for the student. Students who completed this course were successful in obtaining significant gains from the first assessment to the second assessment.

With individual, descriptive narratives accomplished, the next phase in the second course should entail opportunities to interact synchronously and record those experiences. This proposal is always problematic for the mindset of DE students. With busy schedules of work, school and family, the DE student often interprets an online language course as "self-paced." Coordinating with other students' busy lives can create stress and frustration.

Retaining DE Students

This study's semester pacing for DE assignments was done in ten-day increments. The rationale was to provide the most flexible schedule for DE learners. Students who withdrew from the course were those that fell behind in submitting work. They often attempted to complete the entire week's work in one to three hours prior to the due date. The first recommendation is to better pace the weekly work. One of the limitations of the study was uneven groups. This proved problematic when employing statistical analyses in terms of effect size or strength of association. This is perhaps the reason many studies collect data over several semesters (Blake, 2008; Despain, 2003).

The DE sections had high attrition rates. DE attrition rates of 51.8% (Required Speaking) and 79% (Optional Speaking) compared to the f2f sections of 11% (Required Speaking) and 29.6% (Optional Speaking). Pacing student work may help these students comply with assignments and resist the temptation to complete work in one hour the night before the due date. The recommendation is to create a week's schedule similar to those attending a four-day a week class with assignments due Monday through Thursday as those who attend a f2f class. The pacing allows for closer monitoring of student work by the student himself as well as the instructor.

A second recommendation is to reduce the course length from 15 weeks. It might improve retention to divide the course in two parts so that those who complete 8 weeks of the first half can decide to continue or discontinue. The course duration and content should determine the credit hour component.

The Use of Digital Materials

There are so many activities online and offline that accompany a publisher produced package/bundle. Strategically selecting the activities that provide "a bang for your buck" is important. It is important to have on-going discussions with students to evaluate their learning, reaction to course materials, and feedback to activities to inform future courses. Because only a few students will do *suggested* activities that are not tied to grades or extra credit, it is important to create, select and organize the sequence of activities for the DE student and have each tied to a part of the final grade. Alternatively, student produced projects are often more inviting and motivating for the student than a series of exams or 50 activities.

LIMITATIONS OF THE STUDY

Course Participants

The first and foremost important limitation to the study is that comparison of DE and f2f student oral performances represent a snapshot in time. The study began with 108 participants; 27 each section. The attrition rates for the distance education sections were very high: retaining only 13 and eight participants in each DE group. Although post hoc tests were performed where statistical significance was found in this study, small numbers may produce false-positive results. Why did students withdraw at higher rates from the DE sections is an important question to research further.

The Study Design

This was a quasi-experimental design with the primary goal of investigating end of semester measures in speaking as exploring the use of speaking tasks to promote developing speaking skills. It was difficult to anticipate design issues in advance with

having so few empirical studies to draw upon. It would have been useful to monitor a set of speaking practice activities performed live from each section to compare with practice WAV files. There is not much indication that many of the optional sections completed the basic speaking tasks/activities from the textbook. A failure to track confidence levels and beliefs at the beginning and middle of the course is a limitation to this study.

The Instructor/Researcher

Being a participant observer in this quasi-experimental study was not always easy. I wanted all the students to participate fully and was discouraged and disappointed when students withdrew from the course (for a myriad of reasons). It was difficult to "act normal" and not overly intervene[13] with poorly performing students. I kept a teaching journal to separate my own personal feelings from my "teacher's" professional duties and obligations. Balancing the role of researcher and classroom instructor provided two views of evaluating not only the students but of myself, as teacher. I hope my role did not interfere with student learning outcomes and performances as I hoped to gather them without bias. The instructor/researcher role, as participant-observer, always brings into question any unintended influence which may limit the study.

RECOMMENDATIONS FOR FUTURE STUDIES

There are several future studies that are directly related to the questions raised in this study. There are issues concerning attrition and speaking test strategies. Attrition issues might include not only outside issues of work and family or academic preparation but may also include environmental issues within the class context: bonding and trust.

[13] I use overly *intervene* because students that need help sometimes don't want it. I feel pedagogical intervention is imperative and professional.

There are also future studies that build on the next level of classroom teaching and learning that are worth mentioning. I will address both of these issues in more detail in the following section.

Attrition

There are two observations that were made about the course completers: students in the speaking classes had higher retention rates than in the optional speaking classes and students in the f2f classes had higher retention rates than those in the DE classes. Although much is known about the students' learning and social background profiles, it is unknown what may lead students to withdraw from the course. Did they withdraw due to outside pressures such as work and family, or because they were unable to keep pace with the course? Alternately, did they drop the course for academic reasons? Because most students that completed the course had previous language learning experience, a detailed analysis of non-completers' previous experience with language learning is also warranted.

It is not uncommon to have family members in community college classes. This study included two f2f sections with family pairs or groups. The f2f with required speaking had a group of a brother and two sisters and a pair of sisters. The f2f with optional speaking had a mother-daughter pair. Did having family members in the class create an environment of trust? Is bonding - leading to trust - among classmates a feature important for students to take risks in TL language output practice and use? How might *bonding* contribute to class retention?

Strategies for Speaking Tests

This study did not track the learners' test taking behavior. Unknown variables include: test item order of completion, time spent planning for each item, and planning strategies employed. There is evidence, and fascinating data at that, that all the non-heritage students in this study wrote notes and sentences on their test booklets. This was in contrast to all the heritage learners, who did not make notes for themselves, regardless of their language abilities. An entire study is warranted to discover what students do to prepare to speak when given planning time. How does their writing to speak match up with their actual audio files? This study did not compare the midterm booklet writings to the final writings. Do learners' test-taking strategies change with experience and if so, how? Successful strategies may help students to practice and perform better as demands for classroom TL use increase with each lesson and semester. Future research projects will need to seek answers to all of these questions.

Successive Level of Spanish

With the next level of beginning Spanish, students have a set (albeit small) of vocabulary and basic structures from which to pull for communicative purposes. How will students perform with recorded paired tasks on practice activities, repeated activities and assessment items? Today's online tools provide tracking for such synchronous communication (see Illustration 10).

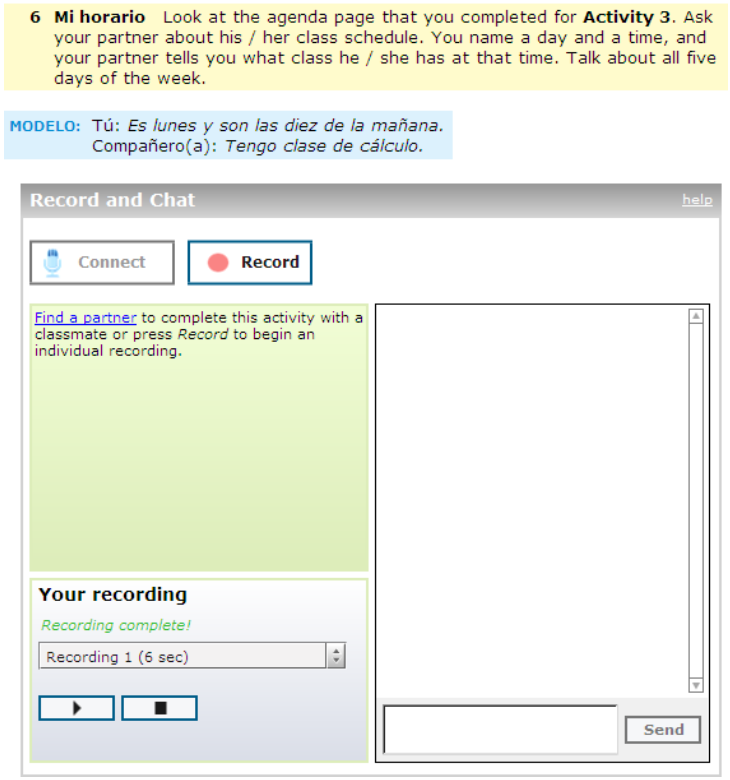


Illustration 10: Online Screen Shot of Interactive Speaking Activity

The students in pairs or small groups can sit online together and record the performance. The recording can be erased and re-recorded. The "send" button holds the recording in a student's individual gradebook for feedback. Instructors can listen to the entire performance and provide two students at one time. All speaking can now be monitored and tracked.

Another related topic is working in 3-D virtual worlds. With synchronous audio capabilities in worlds such as Second Life (sl), how can SL speaking tasks be designed for second semester beginning students and how can the performances compare to the web-based performances (see Illustration 11)?

Illustration 11: A Virtual Fieldtrip to Mexico: Exploring, Speaking & Doing

In this illustration, an avatar teacher, rides with her students (one in photo) through the Mexican Tourist virtual island in Second Life. The student controls the movement of the avatar and with a headphone/microphone set can communicate with other avatars in the immediate proximity. Are there advantages to speaking in 3-D versus web-based pages?

Distance Education: A Question of Engagement

How can we measure engagement? How can we encourage compliance with live attendance? When Wimba live classrooms are archived for later viewing, does this motivate students to participate? Would having an un-archived class, such as in Second Life, motivate students to participate?

CLOSING REMARKS

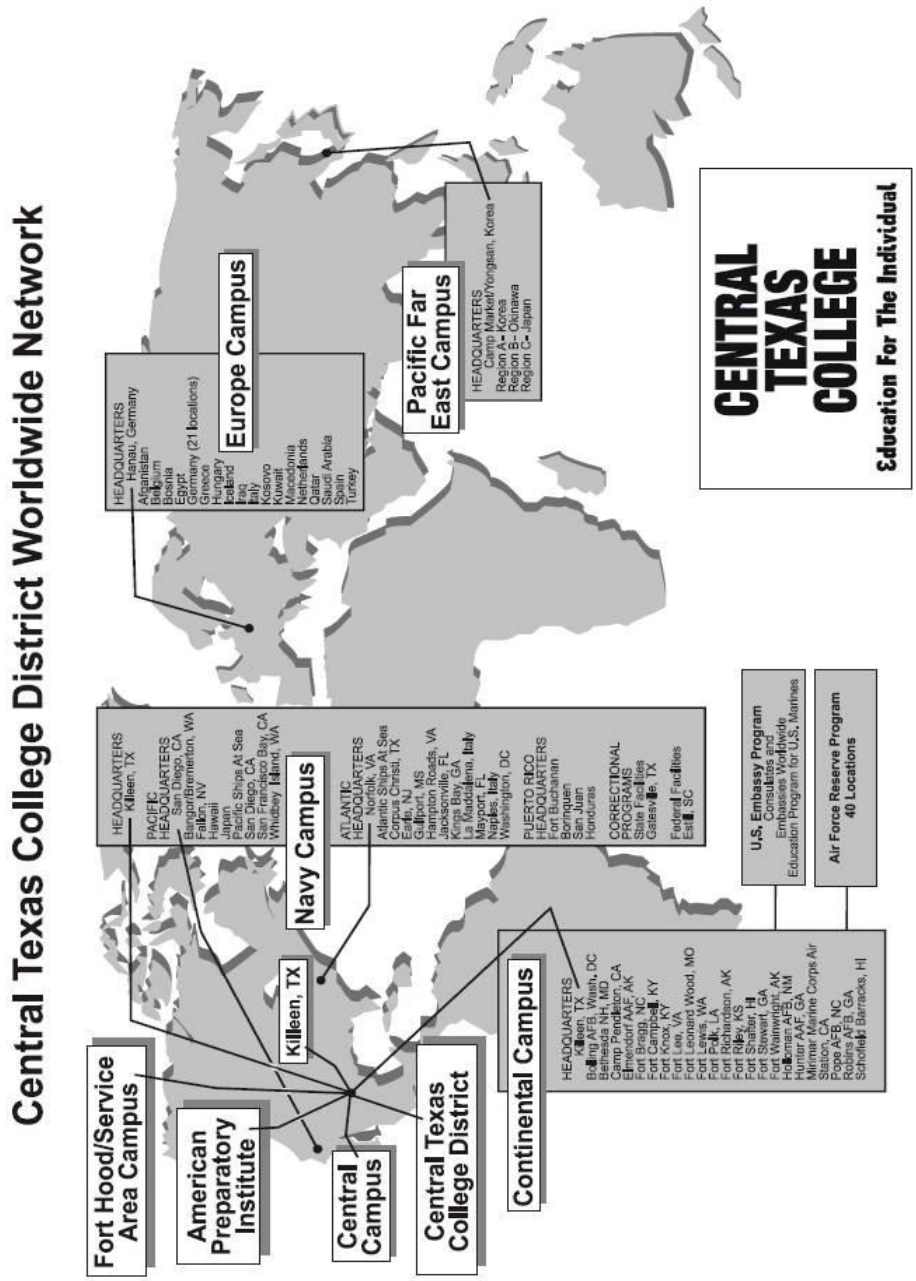
Distance education foreign language courses are now being offered almost as regularly as the classroom-based courses. It is a challenge to participate in the teaching of language courses at a distance. This study was not only experimental but practical with real learners who were motivated by receiving academic credits and grades, and improving language abilities. The results of this study indicate that participants can

obtain equal speaking levels. For those who participated in the creation of WAV files, the performance levels were higher.

Instructors of distance education foreign language courses must provide opportunities (albeit tied to grades) for students to engage in spoken activities. Tests should be transparent to reflect course objectives and realistic practice. Very few studies have posed the questions of how speaking can develop online at a distance and if the performances are equal to classroom participants. This is one of the challenging skills to address and assess. There is the need for further research to add to our understanding of online distance education speaking processes and development.

Appendix A

Central Texas College Map



Appendix B

My Personal Learning Biography

PART 1: Chronological overview of my language learning experience

Year/Data	Languages I grew up with, Language areas I have lived in, Schools/Courses (add details of length, hours), Practical language use at work, training or with acquaintances, on trips etc. Learning experiences, learning progress.

PART 2: Self-assessment & Objectives

If this is a ***skill*** you have already, check it in the (S) column.
☑ I can do this under normal circumstances
☑☑ I can do this easily

If this is an ***objective*** for you, mark it in the (O) column.
☒ This is an objective for me.

Rank your top objectives in the (R) column.
Number 1 is the most important to you, followed by number 2 etc.

	S	O	R
LA1) I can understand when someone speaks very slowly and articulates carefully, with long pauses for me to assimilate meaning.			
I can understand simple directions – how to get from X to Y			
I can understand questions & instructions addressed carefully & slowly to me and follow short, simple directions.			
I can understand numbers, prices and times.			
RA1) I can understand information about people in the newspaper.			
I can locate a concert or a film on calendars of public events, posters & identify where, when & what time it will take place.			
I can understand a questionnaire (hotel registration) to give the most important information about myself – name, DOB, nationality.			
I can understand words/phrases on signs encountered in everyday life. (station, parking, no parking, no smoking, keep left etc.)			
I can understand orders on a computer program such as print, save & copy.			
I can follow short simple written directions.			
I can understand short messages on postcards and email.			
I can understand simple messages by friends such as “back at 4:00.”			
SIA1) I can introduce myself and use basic greeting/leave taking expressions.			
I can ask and answer simple questions, initiate and respond to simple statements in			

the areas of immediate need or on very familiar topics.			
I can make myself understood in a simple way but I am dependent on my partner being prepared to repeat more slowly and rephrase what I say and to help me to say what I want.			
I can make simple purchases where pointing and other gestures can support what I say.			
I can handle saying numbers, quantities, cost and time.			
I can ask people for things and give people things.			
I can ask people questions about where they live, people they know, things they have, etc. and answer such questions addressed to me provided they are articulated slowly and clearly.			
I can indicate time by such phrases as “next week,” “last Friday,” ”in November,” ”three o’clock.”			
SPA1) I can talk about myself and give personal information (address, phone, age, family, hobbies, nationality etc.)			
I can describe where I live.			
SPA2) I can describe my family and other people.			
I can give short basic descriptions of events.			
I can describe my hobbies and interests in a simple way.			
I can describe past activities and personal experiences (last weekend…)			
I can describe my educational background, my present/most recent job.			
LSA1) I can say when I don’t understand.			
I can very simply ask someone to repeat what they said.			
I can very simply ask someone to speak more slowly.			
WA1) I can fill out a questionnaire with my personal data.			
I can write a greeting card or short email.			
I can write a short message to tell someone where I am or where to meet.			
I can write sentences and phrases about myself – where I live and what I do.			

List what you <u>most</u> want to be able to do by the end of the first semester of beginning Spanish.

I want to be able to…….

What are you reasons for taking this course?

Appendix C

My Technology and Personal Profile

Overview of my technological experience

- Please describe the type of computer to which you have regular access.
- Do you have a microphone, headphone and/or working speakers?
- What type of software do you have and know how to use?
- Do you have internet access? Describe your ISP and connection.
- Do you email?
- Have you used discussion boards before?
- Have you chatted online before? (specify voice chat or written chat)
- Have you ever taken an online class?

Personal Profile

- How many classes do you have this semester? Online/In class
- How many hours a week do you work?
- With whom do you live? How old are they?
- What time zone do you currently live in?
- Do you plan to move during the semester?

Appendix D

Speaking Assessment Rater Forms

Speaking Assessment – Rater Forms Participant #_____ Rater#_____

Comprehensibility (Adapted from various sources. Borrás & Lafayette, 1994)
General Definition:

0	1	2	3	4	5
No utterances rendered	Virtually no relevant information was conveyed by the speaker.	Very little information was conveyed that could be understood by a sympathetic native speaker listener. (i.e. a word or two)	Some information was conveyed that could be understood by a native sympathetic native speaker listener. (i.e. phrases)	All information conveyed could be understood by a native sympathetic speaker/listener	All information could be understood by an unsympathetic native speaker/listener

Fluency
General Definition: The overall smoothness, continuity, and naturalness of the subject's speech as opposed to pauses for rephrasing sentences, groping for words and so forth.

0	1	2	3	4	5
No utterances rendered	Utterances halting and fragmentary that communication is virtually impossible.	Utterances very slow, uneven, and often incomplete.	Utterances fairly slow, hesitant, and uneven.	Utterances produced at a reasonable speed though with occasional hesitancies.	Utterances fast and seem effortless

Vocabulary
General Definition: The amount and range of vocabulary usage as deemed appropriate to respond to the prompt.

0	1	2	3	4	5
No utterances rendered	Very limited vocabulary, frequent use of inappropriate terms	Limited vocabulary, some inappropriate terms used	Moderate vocabulary, a few inappropriate terms	Good vocabulary usage, almost no inappropriate terms	Good extensive vocabulary, all terms appropriate to topic

Accuracy (adapted from Borrás & Lafayette, 1994)
General Definition: The grammatical/morpho-lexical correctness of a speech sample.

0	1	2	3	4	5
No utterances rendered.	No utterances rendered correctly.	Structure of very few utterances rendered correctly.	Some utterances rendered correctly, but with many structural problems.	Many utterances rendered correctly, but few problems remain with structures.	Most utterances rendered correctly; no problems with structures.

Appendix E

Final: Self-Assessment and Reflection

Please rate yourself 1 through 4 on your performance for each of the speaking tasks you completed for your speaking final. (1 is unconfident and 4 is very confident) Follow your self-rating with any comment you would like to make about your performance.

1. My Introduction Task 1 2 3 4
My Comment:

2. My Weekend Activities Task 1 2 3 4
My Comment:

3. My Family Member Description Task 1 2 3 4
My Comment:

4. Schedule Description Task 1 2 3 4
My Comment:

5. Activities Description & Likes Task 1 2 3 4
My Comment:

6. Describing How I Feel Task 1 2 3 4
My Comment:

Please list the activities you believe were the most helpful in developing your speaking skills this semester.

Please list the activities you believe were the least helpful in developing your speaking skills this semester.

Appendix F

The Final Speaking Assessment

Name: _____
Date:_____

START TIME: _____

Speaking Assessment Item 1

You happen to run into your Spanish Instructor who does not speak any English. In Spanish, greet her and introduce her to a family member.

Save your voice recording before doing the next speaking item.
Save as **last name** + **first name initial** + **1**. (Example: SMITHB1.wav)

Speaking Assessment Item 2

A Mexican friend is learning about U.S. customs. In Spanish, describe how you (alone or with family or friends) usually spend the weekends. Also include what you are planning to do this upcoming weekend.

Save your voice recording before doing the next speaking item.
Save **as last name** + **first name initial** + **2**.

Speaking Assessment Item 3 

One of your Dominican friends really wants you to practice speaking Spanish more often. As you meet for coffee, you discuss your families in Spanish. Choose one family member to describe in detail. Include the following: name, age, physical description, where he/she lives, employment status, and something about his/her personality.

Save your voice recording before doing the next speaking item.
Save as **last name** + **first name initial** + **3**.

Speaking Assessment Item 4

You are hosting a high school exchange student from Ecuador for one week. She has never been in the U.S. before and does not speak much English. You need to use your Spanish to explain her week's schedule to her. Today is Sunday. Use the following schedule. Try to make at least 5 sentences.

Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Arrive and meet host families	School 8 AM – 3:30 PM 	School 8 AM – 3:30 PM	School 8 AM – 3:30 PM	7:00 AM  Travel to the Capital 	School 8 AM – 3:30 PM	Return Home 
 		Movies 7:00 PM 		Visit History Museum  Return 5:30 PM	Football Game 6:30 PM  Dance 9:00 PM 	Airport 1:00 PM 

Save your voice recording before doing the next speaking item.
Save as **last name** + **first name initial** + **4**.

Speaking Assessment Item 5

Using the following pictures, describe what everyone is doing and say whether or not you like to do the same thing.

a) 

b) 

c) 

d) 

Save your voice recording before doing the next speaking item.
Save as **last name** + **first name initial** + **5**.

Speaking Assessment Item 6

You have entered into a Yahoo Spanish chat room. There are many topics crisscrossing the conversation page. Sometimes the questions and comments appear quite random. Here are several questions directed to you. Since you have a microphone, you have decided it is faster to voice your responses in Spanish instead of type your responses.

a) Chica_amiga asks you to describe how you feel after 4 hours of sleep.

b) Miguel_el_guapo wants to know how you are before an important exam or work deadline.

c) Catracho_cuatro asks how you would feel after losing your backpack with your car keys and wallet.

Save your voice recording.
Save as **last name** + **first name initial** + **6**.
This is the end of your speaking assessment.

FINISH TIME: _____
(Hand this to your proctor.)

After handing this to your proctor, go to the desktop and submit your 6 speaking files to the BLACKBOARD course digital drop box.

Appendix G

Example of Lexical Range/Complexity Analysis

Lexical Complexity Measure: This is a measure of lexical diversity or variety or range. In this study, it is the simplest of all formulas: the number of word types to the number of word tokens (Ortega, 2009: 124).

Embedded forms: Es (She is) = subject pronoun/verb, two types/two tokens
Viva (live!) = verb, one type/one token

Separated form: Ella trabaja. = subject pronoun/verb, two types/two tokens

Here are two full examples from the data:

Participant 52 Performance Final #3

> Mi esposa Janet tiene "twenty-two."
>
> Es guapa. Viva en Filipinas. Ella trabaja.

Lexical Types with (Lexical Tokens):

1. Possessive Pronoun (1)
2. Noun (2)
3. Verb (4)
4. Adjective (1)
5. Preposition (1)
6. Subject Pronoun (2)

Formula: 6 / 11 = .55

Participant 53 Performance Final #3

> Yo tengo un hermano. Su nombre es James. Y tiene, él tiene treinta y nueve años. El es muy alto y vive en Mississippi. El es muy cómico y trabajador.

Lexical Types with (Lexical Tokens):

1. Possessive Pronoun (1)
2. Noun (5)
3. Verb (7)
4. Adjective (3)
5. Preposition (1)
6. Subject Pronoun (6)
7. Quantifier (1)
8. Article (1)
9. Intensifier (2)
10. Connector (3)

Formula: 10 / 30 = .33

Appendix H

Textbook & Student Activities Manual: Skills & Feedback

Chapter 1 p. 8 Vocabulario 1 VV (Vocabulary Video - dialog) R (Repaso)		
Ch 1 p. 9 Vocabulario 1 Presentation Actividad 1 MP (más práctica) MiniP (mini prueba)	**Type of Activity** Multiple choice Crossword Multiple choice	**Type of Feedback** Graded Graded Graded
Ch 1 p. 11 Vocabulario 2 Presentation VV Repaso (review)		
Ch 1 p. 12 Vocabulario 2 Actividad 2 Activiadad 3 MP MiniP	 Speaking/writing Speaking Listening/Speaking Writing	 Instructor graded Non graded Non graded Graded
Ch 1 p. 13 Vocabulario 3 Presentation VV Repaso (review)		
Ch 1 p. 14 Vocabulario 3 Actividad 4 Actividad 5 Actividad 6 MP MiniP	 Dialog writing Writing Writing Listening – match Writing Listening/Writing	 Instructor graded Instructor graded Instructor graded Non graded Graded Non graded
Ch 1 **Student Activities Manual** ¡Imagínate! Vocabulario	**Workbook** 1, 2, 3, 4, 5, 6	**Lab Manual** Pronunciación 1, 2,3,4
Ch 1 p. 19 Gramática 1 Presentation Repaso (review) Actividad 1	 Listening/Writing	 Graded
Ch 1 p. 20 Gram. 1 Actividad 2 Actividad 3 Actividad 4 MP	 Writing Writing Speaking m/c – game	 Graded Graded Non graded Graded

MiniP	writing & m/c	Graded
Ch 1 p. 22 Gram. 2 Presentation VT (video tutorial) R Actividad 5	 Writing	 Graded
Ch 1 p. 23 Gram. 2 Actividad 6 Activiadad 7 Actividad 8 MP MiniP	 Writing Matching Speaking Matching – m/c m/c Listening – Writing Writing	 Instructor graded Graded Non graded Graded Graded Instructor graded Graded
Ch 1 p. 24 Gram. 3 Presentation R Actividad 9 Actividad 10 Actividad 11 MP MiniP	 Writing Writing – Speaking Speaking Matching Writing Writing	 Graded Non graded Non graded Graded Instructor graded Graded
Ch 1 p 27 Gram. 4 Presentation R VT Actividad 12 Actividad 13 Actividad 14 MP MiniP	 Writing Speaking Listen & write Matching Listening & writing m/c - writing	 Instructor graded Non graded Graded Graded Instructor graded Graded
Ch 1 – **Student Activity Manual** ¡Prepárate! Gramática	**Workbook** 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	**Lab Manual** Comprensión 5, 6, 7, 8, 9
Student Activity Manual A Leer	**Workbook** Reading: 18, 19, 20	
Escribir	Writing: 21, 22	

Appendix I

UT IRB Approval Letter

OFFICE OF RESEARCH SUPPORT & COMPLIANCE

THE UNIVERSITY OF TEXAS AT AUSTIN

P.O. Box 7426, Austin, Texas 78713 (512) 471-8871 - FAX (512 471-8873)
North Office Building A, Suite 5.200 (Mail code A3200)

FWA # 00002030

Date: **09/21/06**

PI(s): **Lisa M Volle** Department & Mail Code: ***EMPLOYEE TERMINATED*** **Y0002**

Dear: **Lisa M Volle**
IRB APPROVAL – IRB Protocol # **2006-08-0054**

Title: **Teacher Action Research: A Comparative Study of the Speaking Performance of On-line and In-class Sections of**

In accordance with Federal Regulations for review of research protocols, the Institutional Review Board has reviewed the above referenced protocol and found that it met approval under an Expedited category for the following period of time: **09/21/2006 - 09/20/2007**

Expedited category of approval:

___**(1)** Clinical studies of drugs and medical devices only when condition (a) or (b) is met. (a) Research on drugs for which an investigational new drug application (21 CFR Part 312) is not required. (Note: Research on marketed drugs that significantly increases the risks or decreases the acceptability of the risks associated with the use of the product is not eligible for expedited review). (b) Research on medical devices for which (i) an investigational device exemption application (21 CFR Part 812) is not required; or (ii) the medical device is cleared/approved for marketing and the medical device is being used in accordance with its cleared/approved labeling.

___**(2)** Collection of blood samples by finger stick, heel stick, ear stick, or venipuncture as follows: (a) from healthy, non-pregnant adults who weigh at least 110 pounds. For these subjects, the amounts drawn may not exceed 550 ml in an 8 week period and collection may not occur more frequently than 2 times per week; or (b) from other adults and children2, considering the age, weight, and health of the subjects, the collection procedure, the amount of blood to be collected, and the frequency with which it will be collected. For these subjects, the amount drawn may not exceed the lesser of 50 ml or 3 ml per kg in an 8 week period and collection may not occur more frequently than 2 times per week.

___**(3)** Prospective collection of biological specimens for research purposes by Non-invasive means. Examples:

(a) hair and nail clippings in a non-disfiguring manner;
(b) deciduous teeth at time of exfoliation or if routine patient care indicates a need for extraction;
(c) permanent teeth if routine patient care indicates a need for extraction;
(d) excreta and external secretions (including sweat);
(e) uncannulated saliva collected either in an un-stimulated fashion or stimulated by chewing gumbase or wax or by applying a dilute citric solution to the tongue;
(f) placenta removed at delivery;
(g) amniotic fluid obtained at the time of rupture of the membrane prior to or during labor;
(h) supra- and subgingival dental plaque and calculus, provided the collection procedure is not more invasive than routine prophylactic scaling of the teeth and the Process is accomplished in accordance with accepted prophylactic techniques;
(i) mucosal and skin cells collected by buccal scraping or swab, skin swab, or mouth washings;
(j) sputum collected after saline mist nebulization.

☐ **(4)** Collection of data through noninvasive procedures (not involving general anesthesia or sedation) routinely employed in clinical practice, excluding procedures involving x-rays or microwaves. Where medical devices are employed, they must be cleared/approved for marketing. (Studies intended to evaluate the safety and effectiveness of the medical device are not generally eligible for expedited review, including studies of cleared medical devices for new indications). Examples:

(a) physical sensors that are applied either to the surface of the body or at a distance and do not involve input of significant amounts of energy into the subject or an invasion of the subject's privacy;
(b) weighing or testing sensory acuity;
(c) magnetic resonance imaging;
(d) electrocardiography, electroencephalography, thermography, detection of naturally occurring radioactivity, electroretinography, ultrasound, diagnostic infrared imaging, doppler blood flow, and echocardiography;
(e) moderate exercise, muscular strength testing, body composition assessment, and flexibility testing where appropriate given the age, weight, and health of the individual.

___ **(5)** Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected solely for non-research purposes (such as medical treatment or diagnosis). (NOTE: Some research in this category may be exempt from the HHS regulations for the protection of human subjects. 45 CFR 46.101(b)(4). This listing refers only to research that is not exempt).

_X_ **(6)** Collection of data from voice, video, digital, or image recordings made for research purposes.

_X_ **(7)** Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. (NOTE: Some research in this category may be exempt from the HHS regulations for the protection of human subjects. 45 CFR 46.101(b)(2) and (b)(3). This listing refers only to research that is not exempt).

_X_ **Please use the attached approved informed consent**

____You have been granted Waiver of Documentation of Consent
According to 45 CFR 46.117, an IRB may waive the requirement for the investigator to obtain a signed consent form for some or all subjects if it finds either:

___ The research presents no more than minimal risk

AND

___ The research involves procedures that do not require written consent when performed outside of a research setting

<OR>

___ The principal risks are those associated with a breach of confidentiality concerning the subject's participation in the research **AND**

___ The consent document is the only record linking the subject with the research

AND

___This study is not FDA regulated (45 CFR 46.117)

AND

___Each participant will be asked whether the participant wishes documentation linking the participant with the research, and the participants wishes will govern.

____ You have been granted Waiver of Informed Consent
According to 45 CFR 46.116(d), an IRB may waive or alter some or all of the requirements for Informed consent if:

___ The research presents no more than minimal risk to subjects;
___ The waiver will not adversely affect the rights and welfare of subjects;
___ The research could not practicably be carried out without the waiver; and
___ Whenever appropriate, the subjects will be provided with additional pertinent information they have participated in the study.

___ This study is not FDA regulated (45 CFR 46.117)

RESPONSIBILITIES OF PRINCIPAL INVESTIGATOR FOR ONGOING PROTOCOLS:

(1) Report **immediately** to the IRB any unanticipated problems.

(2) Proposed changes in approved research during the period for which IRB approval cannot be initiated without IRB review and approval, except when necessary to eliminate apparent immediate hazards to the participant. Changes in approved research initiated without IRB review and approval initiated to eliminate apparent immediate hazards to the participant must be promptly reported to the IRB, and reviewed under the unanticipated problems policy to determine whether the change was consistent with ensuring the participants continued welfare.

(3) Report any significant findings that become known in the course of the research that might affect the willingness of subjects to continue to take part.

(4) Insure that only persons formally approved by the IRB enroll subjects.

(5) Use **only** a currently approved consent form (remember approval periods are for 12 months or less).

(6) **Protect the confidentiality of all persons and personally identifiable data, and train your staff and collaborators on policies and procedures for ensuring the privacy and confidentiality of participants and information.**

(7) Submit for review and approval by the IRB all modifications to the protocol or consent form(s) prior to the implementation of the change.

(8) Submit a **Continuing Review Report** for continuing review by the IRB. Federal regulations require **IRB review of on-going projects no less than once a year** (a Continuing Review Report form and a reminder letter will be sent to you 2 months before your expiration date). Please note however, that if you do not receive a reminder from this office about your upcoming continuing review, it is the primary responsibility of the PI not to exceed the expiration date in collection of any information. Finally, it is the responsibility of the PI to submit the Continuing Review Report before the expiration period.

(9) Notify the IRB when the study has been completed and complete the Final Report Form.

(10) Please help us help you by including the above protocol number on all future correspondence relating to this protocol.

Thank you for your help in this matter.

Sincerely,

Lisa Leiden, Ph.D., Chair
Institutional Review Board

Protocol # Approval dates: - 2006-08-0054 09/21/2006 09/20/2007

Appendix J

CTC IRB Approval Letter

CENTRAL TEXAS COLLEGE
Education For The Individual

Mr. Don Mikles
Dean, Central Campus
254/526-1116
Don.Mikles@ctcd.edu

August 14, 2006

Dr. Lisa Leiden, Ph.D.
Director, Office of Research Support and Compliance
P.O. Box 7426 Campus Mail
Austin, TX 78713
Lisa.leiden@mail.utexas.edu

Dear Dr. Leiden:

The purpose of this letter is to grant Lisa M. Volle, a graduate student at the University of Texas at Austin permission to conduct research at Central Texas College. The project, Teacher Action Research: A Comparative Study of the Speaking Performance of On-line and In-class Sections of a First Semester Foreign Language Spanish Class, entails collecting oral speech samples of students from various Distance Education Testing locations and from the Language classroom at the main campus by the Instructor/Researcher for four sections of first semester Spanish. Each section consists of 27 enrolled students. All students are expected to be adults over age 18. The purpose of the proposed research is to inform the language program about the quantity and quality of speaking performance in the various learning contexts. Central Texas College was selected because this is an important skill area in language learning, and this institution is interested in improving learning and instruction for our various learning contexts here and at a distance. Lisa Volle has been a full-time faculty member at Central Texas College since 1987 and is actively involved in critical issues with technology and learning. She will be sharing the results of her findings on this campus and with our distance education department in writing by including her findings and recommendations. I, Don Mikles, Dean of Central Campus, do hereby grant permission for Lisa Volle to conduct "Teacher Action Research: A Comparative Study of the Speaking Performance of On-line and In-class Sections of a First Semester Foreign Language Spanish Class" at Central Texas College.

Sincerely,

P. O. Box 1800 Killeen, Texas 76540-1800
and over 100 other locations worldwide!

Bibliography

Abrams, Z. (2003). The effect of synchronous and Asynchronous CMC on oral performance in German. *The Modern Language Journal*, 87(2), 157-167.

Abrioux , D. A. M. X. (1989). Computer-Assisted Language Learning and Distance Education. *Journal of Distance Education, 4*, 20-35.

ACTFL Proficiency Guidelines, 1995. (retrieved from the WWW, June, 10, 2003.)

ACTFL Standards for Foreign Language Learning in the 21st Century, (2006). Lawrence, KS: Allen Press, Inc.

Adair-Hauck, B., Willingham-McLain, L., & Earnest-Youngs, B. (1999). Evaluating the integration of technology and second language learning. *CALICO Journal,* 17(2), 269-306.Aemmer, M. (2006). Top 10 Biggest Colleges: For-profit schools and community colleges rising, Retrieved on December 10, 2006 from http://www.skyscrapercity.com/showthread.php?t=367165.

Ahmadian , M. J., Tavakoli, M., Ketabi, S. & Kasaian, Z. (2010). On the benefits of careful within-task planning and task repetition in EFL classrooms. *English Language Teaching*, 3(1), 155-158.

Aitsiselmi, F. (1999). Second language acquisition through email interaction. *ReCALL*, 11(2), 4–11.

Ariza, E. N., & Hancock, S. (2003). Second language acquisition theories as a framework for creating distance learning courses. *International Review of Research in Open and Distance Learning*, 4(2), Retrieved April 3, 2005 from http://www.irrodl.org/content/v2004.2002/ariza-hancock.html.

Bachman, L. F. (1990). Fundamental considerations in language testing. Oxford: Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice. Oxford: Oxford University Press.

Bachman, L. F. (1991). What does language testing have to offer?. *TESOL Quarterly.* 25(4).

Bailey, K. M. (1998). *Learning about Language Assessment*. Boston: Newberry House Heinle & Heinle.

Bailey, K. M. (1996). Working for washback: a review of the washback concept in language testing. *Language Testing*, 13, 257 –279.

Bailey, K. M. (1992). The processes of innovation in language teacher development: What, why and how teachers change. In J. Flowerdew, M. Brock, and S. Hsia (Eds.), *Perspectives on Second Language Teacher Education*. Hong Kong: City Polytechnic of Hong Kong, pp. 253-282. Retrieved on June, 10, 2006 from http://www.kathleenmbailey.com/2_pubs_articles_pdf.html.

Blake, R. (2008). Brave New Digital Classroom: Technology and Foreign Language Learning. Washington, D.C: Georgetown University Press.

Blake, R. (2000). Computer mediated communication: A Window on L2 Spanish Interlanguage. Language Learning & Technology, 4, 120-136.

Blake, R., & Delforge, A. (2007). Online Language Learning: The Case of Spanish Without Walls. In B. Lafford & R. Salaberry (Eds.), The art of teaching Spanish: Second language acquisition from research to praxis (pp. 127-147). Georgetown: Georgetown University Press.

Blake, R., Wilson, N. L., Cetto, M., & Pardo-Ballester, C. (2008). Measuring oral proficiency in distance, face-to-face, hybrid and blended classrooms. *Language Learning & Technology*, 12(3), 114-127.

Borrás, I. (1993). Developing and assessing "Practicing Spoken French": A multimedia program for improving speaking skills. Educational Technology Research and Development, 41(4), 91-103.

Borrás, I, & Lafayette, R. (1994). Effects of Multimedia Courseware Subtitling on the Speaking Performance of College Students of French. The Modern Language Journal, 78, 61-75.

Brushwood, J., & Polmantier, P. (1953). *The effectiveness of the audio laboratory in elementary Modern Language courses.* Columbia, MO: The University of Missouri.

Buka, M., Freeman, M. K., & Locke, W. N. (1962). Language teaming and frequency response. *International Journal of American Linguistics* ,28, 62–79.

Bygate, M. (2001). Effects of Task Repetition on the Structure and Control of Oral Language. In M. Bygate, P. Skehan & M. Swain (Eds.), Researching Pedagogic Tasks (23-48). Harlow, England: Longman.

Cahill, D., & Catanzaro, D. (1997). Teaching first-year Spanish on-line. *CALICO Journal*, 14(2), 97-114.

Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second-language teaching and testing. *Applied Linguistics*, 1(1), 1-47.

Celce-Murcia, M. Dörnyei, Z., and Thurrell, S. (1997). Direct approaches in L2 instruction: A turning point in communicative language teaching?. TESOL Quarterly, 31, 141-152.

Chenoweth, N. A., & Murday, K. (2003). Measuring student learning in an online French course. *CALICO Journal*, 20(2), 284-314.

Chomsky, N. (1970). Methodological Preliminaries. *Aspect of the Theory of Syntax*, 26-40.

Chomsky, N. (1986). *Knowledge of Language*, Praeger.

Chun, D. (1992). The State of the art in teaching pronunciation. In J. Alatis (Ed.), Linguistics and Language Pedagogy: The State of the Art. Gerorgetown University round Table on Languages and Linguistics (pp. 179-193). Washington, DC: Georgetown University Press.

Chun, D. (1994). Using Computer Networking to Facilitate the Acquisition of Interactive Competence. *System*, 22, 17-31.

Chun, D. (2007). Come ride the wave: But where is it taking us? *CALICO*, 24(2): 239-252.

Clamphan, C. & Corson, D. (Eds) (1998). Language Testing and Assessment, Vol. 7 of the *Encyclopedia of Language and Education*. Amsterdam: Kluwer Academic Publishers.

Clifford, R. T. (1998). Mirror, mirror, on the wall: Reflections on computer assisted language learning. *Calico Journal*, 16(1), 1-10.

Cohen, A. D. (1988). *Statistical Power Analysis for the Behavioral Sciences* (Second ed.). Lawrence Erlbaum Associates.

Coniam, D. (1998). The use of speech recognition software as a English language oral assessment instrument: An exploratory study. *CALICO Journal*, 15(4), 7-23.

COOK, V. (2001). *SECOND LANGUAGE LEARNING AND LANGUAGE TEACHING* (3RD EDITION). LONDON: ARNOLD, CO-PUBLISHED BY OXFORD UNIVERSITY PRESS, NEW YORK.

Crookes, G. (1989). Planning and Interlanguage Variation. Studies in Second Language Acquisition, 11, 367-383.

Crookes, G. (1990). The utterance, and other basic units for second language discourse analysis, Applied Linguistics, 11(2), 183-199.

Derwing, T., Munro, M., & Carbanaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly*, 24 (4), 592-603.

Despain, J. S. (2003). Achievement and attrition rate differences between traditional and internet-based beginning Spanish courses. *Foreign Language Annals*, 36(2), 243-257.

Doughty, C. J. & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology*, 7(3), 50-80.

Doughty, C. & Williams, J. (1998). *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press

Egan, K. B. (1999). Speaking: A critical skill and challenge. *CALICO Journal*, 16(3), 277-293.

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

Ellis, R. (1987). Interlanguage Variability to Narrative Discourse: Style-Shifting in the Use of the Past Tense. *Studies in Second Language Acquisition*, 9, 1-20.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research*, 4 (3), 193-220.

Ellis, R. (2003). *Task-based learning and teaching*. Oxford: Oxford University Press.

Ellis, R. & Barkhuizen, G. (2005). *Analyzing Learner Language*. Oxford: Oxford University Press.

Ellis, R. & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis, (Ed), *Planning and task performance in second language*, 167-192. Amsterdam: John Benjamins Publishing.

European Learning Portfolio (2000). Retrieved on May 12, 2005 from http://www.coe.int/T/DG4/Portfolio/?L=E&M=/main_pages/introduction.html

Foster, P. & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 229-323.

Fotos, J. T. (1955). The Purdue Laboratory Method in Teaching Beginning Classes. *Modern Language Journal*, 39, 141-143.

Fox, M. (1997). Beyond the Technocentric-Developing and Evaluation Content-Driven, Internet-Based Language Acquisition Courses. Computer Assisted Language Learning. 10(5), 443-453.

Fulcher, G. (1998). "Testing Speaking". In Clapham, C. (Ed). *Language Testing and Assessment*. Vol. 7. of the Encyclopaedia of Language and Education, Kluwer Academic Publishers, 75 - 86.

Fulcher, G. and Marquez Reiter, R. (2003). "Task difficulty in speaking tests." *Language Testing*, 20(3), 321 - 344.

Gamlin, G. (2005). Digital voice recordings in online learning environments. *PacCALL Journal* 1(1), 53-62.

Gonzalez-Bueno, M. (1998). The effects of electronic mail on Spanish L2 discourse. Language Learning & Technology, 1, 55-70.

Gonzalez-Edfelt, N. (1990). Oral Interaction and Collaboration at the Computer: Learning English as a Second Language with the Help of Your Peers. Computers in the Schools, 7, 53-90.

Gonzalez Pino, B. (1989). Prochievement Testing of Speaking. Foreign Language Annals, 22(5), 487-496.

Green, A. & Earnest-Youngs, B. (2001). Using the web in elementary French and German courses: Quantitative and qualitative study results. *CALICO Journal*, 19(1), 89-123.

Hampel, R., & Hauck, M. (2004). Towards an effective use of audio conferencing in distance language courses. *Language Learning & Technology, 88*(1), 166-82.

Hanson-Smith, E. (2002). Communities of practice for pre- and in-service teacher education. In *Teacher education in CALL*, P. Hubbard & M. Levy (Eds.), 205-215. Language Learning & Teaching Series, Vol. 14. Philadelphia, PA: John Benjamin.

Hendrickson, J. (1992). Creating listening and speaking prochievement tests. *Hispania*, 75(5), 1326-1331.

Hertel, T. J. (2003). Using an e-mail exchange to promote cultural learning. *Foreign Language Annals*, 36 (3), 386-396.

Ho, Y-K. (2003). Audiotaped dialog journals: an alternative form of speaking practice. *ELT Journal*, 57(3), 269-277.

Holmberg, B., Shelley, M., & White, C. (Eds.). (2004). Distance education and languages: Evolution and change. Clevedon, U.K.: Multilingual Matters.

Hoopingarner, D. (2004). Listening and Speaking Online. Retrieved June 3, 2005 from edu.ocac.gov.tw/discuss/academy/netedu05/html/paper/sw12.pdf

Hurd, S. (2004). Autonomy and the distance language learner in Holmberg, B, M Shelley and C. White (Eds) . *Distance education and languages: evolution and change*. Clevedon, UK: Multilingual Matters

Hurd, S., Beaven, T. & Ortega, A. (2001). Developing autonomy in a distance language learning context: issues and dilemmas for course writers in *System*, 29 (3), 341-355.

Hymes, D. (1974). Foundations in sociolinguistics: An ethnographic approach. Philadelphia: University of Pennsylvania Press.

Iwashita, N., McNamara, T, & Elder. (2001). Can We Predict Task Difficulty in an Oral Proficiency Test? Exploring the Potential of an Information-Processing Approach to Task Design. *Language Learning*, 51(3), 401-436.

Johnson, K. (2003). Designing Language Teaching Tasks. Palgrave Macmillan.

Johnston, J. & Milne, L. (1995). Scaffolding Second Language Communicative Discourse with Teacher-Controlled Multimedia. *Foreign Language Annals*, 28(3), 315- 329.

Katoaka, H.C. (1987). Long distance language teaching: The second year of televised Japanese. *Journal of Educational Techniques and Technologies*, 20(2), 43-50.

Kern, R. (1995). Restructuring Classroom Interaction with Networked Computers: Effects on Quantity and Characteristics of Language Production. The Modern Language Journal, 79(4), 457-476.

Kern, R. G. & Warschauer, M. (2000). Theory and practice of networked-based language teaching. In *Network-based language teaching: Concepts and practice*. M. Warchauer & R. Kern, 1-19. Cambridge: Cambridge University Press.

Koren, S. (1995). Foreign Language Pronunciation Testing: A New Approach. System, 23(3), 387-400.

Kötter, M., Sheilds, L. & Stevens, A. (1999). Real-time audio and email for fluency: promoting distance language learners' aural and oral skills via the internet. *ReCALL* 11(2), 55-60.

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

Krashen, S. (1985). *Language acquisition and language education*. Hayward, CA: Alemany Press.

Lee, L. (2000). Evaluating Intermediate Spanish Students' Speaking Skills through a Taped Test: A Pilot Study. *Hispania*, 83, 127-138.

Lee, L. (2002). Enhancing learners' communication skills through synchronous electronic interaction and task-based instruction. *Foreign Language Annuals*, 35(1), 16-23.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levy, M. (1997). *Computer- assisted language learning: Context and conceptualization*. Oxford, UK: Clarendon.

Levy, S. (2003). *Six factors to consider when planning online distance learning programs in higher education*. Retrieved on April 16, 2003 from http://www.estga.edu/~distance/ojdla/spring61/levy61.htm.

Liaw, M.L. (1997). An analysis of ESL children's verbal interaction during computer book reading. *Computers in the Schools*, 13 (3/4), 55-73.

Little, D. & Perclová, R. (2001). European Language Portfolio: guide for teachers and teacher trainers. Strasbourg: Council of Europe. Retrieved on April 12, 2005 from http://culture.coe.int/portfolio

Liu, M., Moore, Z., Graham, L., & Lee, S. (2002). A look at the research on computer-based technology use in second language learning: A review of the literature from 1990-2000. *Journal of Research on Technology in Education*, 34 (3), 250-273.

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (eds), *Modeling and assessing second language development* , 77-99. Clevedon: Multilingual Matters.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. B. Ginsberg, & C. Kramsch (eds.), *Foreign language research in cross-cultural perspective* , 39-52. Amsterdam: John Benjamins.

Long, M. H. (2006). *Task-based language teaching*. Oxford, England: Basil Blackwell.

Lorge, S. W. (1965). Comments on language laboratory research: A critique. *Modern Language Journal* , 49, 369-370.

Lou, Y., Abrami, P.C., & d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, 71, 449-521.

Lou, Y., Bernard, R. M. & Abrami, P. C. (2006). Media and pedagogy in undergraduate distance education: A theory-based meta-analysis of empirical literature. *ERT&D* , 54(2), 141-176.

Lowe, P. (1983). The ILR oral proficiency interview: Origins, Applications, Pitfalls and Implications. Die Unterrichtspraxis, *16, 230.*

Lowe, P., & Stansfield, C.W. (Eds.). (1988). *Second language proficiency assessment: Current issues*. Englewood Cliffs, NJ: Prentice Hall.

Lumley, T. & O'Sullivan, B. (2005). The Impact of Test Taker Characteristics on Speaking Test Task Performance. *Language Testing*, 22(4), 415-437.

McDonough, S. H. (2001). Promoting self-regulation in Foreign Language Learners. *The Clearing House.* July/August, 323-326.

Mehnert, U. (1998). The Effects of Different Lengths of Time for Planning on Second Language Performance. *Studies in Second Language Acquisition*, 20, 83-108.

Met, M. & Galloway, V. (1996). Research in foreign language curriculum. In *Handbook on research in curriculum*, Ed. P. W. Jackson, 852-590. New York: MacMillan.

Murphy, L. (2005). Attending to form and meaning: the experience of adult distance learners of French, German and Spanish. *Language Teaching Research*, 9(3), 295-317.

Murray, G. L. (1999). Autonomy and language learning in a simulated environment. *System ,* 27, 295-308.

Nunan, D. (1989). Task-based syllabus design: selecting, grading and sequencing tasks. In G. Crookes and S. Cass (eds), *Tasks in Pedagogical Contexts: Integrating theory and practice*. Clevedon: Multilingual Matters, 55-68.

Omaggio, A. H. (2001). Teaching Language in Context. (3rd ed.). Boston: Heinle & Heinle.

Ortega, L. (1995). The effect of planning in L2 Narratives. Honolulu: University of Hawai'i, Second Language Teaching & Curriculum.

Ortega, L. (1999). Planning and Focus on Form in L2 Oral Performance. Studies in Second Language Acquisition, 21, 109-148.

Oxford, R. (1999) Anxiety and the language learner: New insights. In J. Arnold and H. Douglas Brown (eds) *Affect in Language Learning* , 58–67. Cambridge: Cambridge University Press.

Payne, S. (2004). Making the most of synchronous and asynchronous discussion in foreign language instruction. In *The Heinle professional series in language instruction. Vol. I, Teaching with technology*, ed. L. Lomika and J. Cooke-Plagwitz, 79-93. Boston: Heinle.

Payne, J. S, & Whitney, P. J. (2002). Developing L2 Oral Proficiency through Synchronous CMC: Output, Working Memory, and Interlanguage Development. *CALICO Journal*, 20(1), 7-32.

Pethokoukis, J. M. (2002). E-learn and earn. *U.S. News and World Report*, 132(22), 36.

Pica, T. (1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning*, 33 (4), 465-497.

Pica, T. Kanagy, R. & Falodum, J. (1993). Choosing and using communication tasks in second language instruction. In *Tasks and language learning: Integrating theory and practice*, ed. G. Crookes and S. M. Gass, 9- 34. Clevendon: Multilingual Matters.

Pichette, F. (2009). Second Language Anxiety and Distance Learning. *Foreign Language Annals*, 42 (1), 77-93.

Rahimpour, M. (1997). Task complexity and variation in L2 learner's oral discourse. *The University of Queensland Working Papers in Linguistics*, Retrieved on June 25, 2007 from http://espace.library.uq.edu.au/view/UQ:23699.

Riggenbach, H. (1991). Towards an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, 12, 423-441.

Rivers, W. & Temperley, R. S. (1978). *A Practical Guide to the Teaching of English*. New York: Oxford University Press.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, 45, 99-140.

Rosell-Aguilar, F. (2005). Task design for audiographic conferencing: Promoting beginner oral interaction in distance learning. *Computer Assisted Language Learning*, 18(5), 417-442.

Ros i Solé, C. & Mardomingo, R. (2004). Trayectorias: A new model for online task-based learning. *ReCall*, 16(1), 145-157.

Russell, T. L. (1999). *The no significant difference phenomenon: A comparative research annotated bibliography on technology for distance education*. International Distance Education Certification Center.

Salaberry, R. (2001). The use of technology for foreign language learning and teaching: A retrospective. *Modern Language Journal*, 85 (1), 39-56.

Salaberry, M. R. (1997). The development of past tense verbal morphology in classroom L2 Spanish. *Applied Linguistics*, 20, 151-178.

Schick, J. E., & Nelson, P. B. (2001). Language Teacher Education: The Challenge for the Twenty-First Century, *The Clearing House*, 74(6), 301-304

Schmidt, R. & Froda, S. (1986). Developing basic conversational ability in second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition*, 237-322. Rowley, MA: Newberry House.

Shimazu, Y. M. (2005). Language course taught with online supplement material: Is it effective? *Education*, 126(1), 26-36.

Shohamy, E. (1992). Beyond proficiency testing: a diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal*, 76, 513-521.

Shortreed, (1993). Variations in foreign language talk: The effects of task and proficiency. Crookes, G., & Gass, S.M. (Eds.) *Tasks and language learning*. Clevedon, UK: Multilingual Matters.

Shrum, J. L. & Glisan, E. W. (2000). *Teacher's Handbook: Contextualized Language Instruction*, Boston: Heinle

Sisson, C.R. (1970). The effects of delayed comparisons on the language laboratory on phoneme discrimination and pronouncement accuracy. *Language Learning,* 26, 69-88.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2003). FDTL DATA Project: Quantitative Research Methods

Skehan, P. & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49, 93-120.

Sollenberger, H.E. (l978). Development and current use of the FSI Oral Interview test. In J.L.D. Clark, *Direct testing of speaking proficiency: Theory and application* (1-12). Princeton, NJ: Educational Testing Service.

Swaffar, J., Romano, S., Markley, P., & Arens, K. (1998). *Language Learning Online: Theory and Practice in the ESL and L2 Computer Classroom.* Austin, Texas: Labyrinth Publications.

Swain, M. (1985). Communicative Competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in Second Language Acquisition*, 235-253. Rowley, MA: Newberry House.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 50 (1), 158-164.

Swain, M. (1998). Focus on Form through conscious reflection. In G. Cook & B. Seidhofer (Eds.), *Principles and practice in the study of language*, 125-144. Oxford: Oxford University Press.

Tallent-Runnels, M. K. Thomas, J. Lan, W. & Cooper, S. (2006). Teaching Courses Online: A Review of the Research. *Review of Educational Research,* 76 (1), 93-135.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics*, 4, 142-163.

Tarone, E., & Yule, G. (1989). *Focus on the language learner: approaches to identifying and meeting the needs of second language learners*. New York: Oxford University Press.

Taylor, J. C. (1994 – change to 1995). Distance education technologies: The fourth generation. *Australian Journal of Educational Technology*, 11(2), 1-7. Retrieved on April 12, 2005 from http://www.ascilite.org.au/ajet/ajet11/taylor.html.

Thirunarayanan, M. O., & Perez-Prado, A. (2001). Comparing Web-based and classroom-based learning: A quantitative study. *Journal of Research on Computing in Education*, 34 (2), 131-137.

Ting, S. S. C. (1996). Planning time, modality and second language task performance: Accuracy and fluency in the acquisition of Chinese as a second language. *The University of Queensland Working Papers in Language and Linguistics*, 1, 31-64.

Villas Rogers, C. (2002). Tradition and technology in language teaching. *Cyberspace and Foreign Languages: Dimension*, 1(3), 17-35.

Volle, L. (2005). Analyzing Oral Skills in Voice E-Mail and Online Interviews. Language Learning & Technology, 9(3), 146-163.

Wang, Y. (2004). Supporting synchronous distance language learning with desktop videoconferencing. Language Learning and Technology. 8(3), 1-37. http://llt.msu.edu

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal*,13*(*2), 7-26.

White, C. (1995). The metacognative knowledge of distance learners. *Open Learning*, November, 37-46.

White, C. (1999). Expectations and emergent beliefs of self-instructed language learners in *Metacognitive knowledge and beliefs in language learning. System Special Issue 27*, 443-457.

White, C. (2004). Independent Language Learning in Distance Education: Current Issues. In the *Proceedings of the Independent Learning Conference 2003.* Published in September 2004.

White, C. (2005). Contribution of distance education to the development of individual learners. *Distance Education,* 26 (2), 165-181.

Widdowson, H. G. (1979). *Explorations in Applied Linguistics*. Oxford: Oxford University Press.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14, 85-106.

Yuan, F. & Ellis, R. (2003). The Effects of Pre-Task Planning and On-Line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral Production. Applied Linguistics, 24(1), 1-7.

Vita

Lisa Marie Volle was born in Philadelphia, Pennsylvania. She lived in New Orleans and Connecticut before entering Juniata College at the end of her junior year of high school. She majored in international studies and spent her undergraduate junior year abroad in Barcelona, Spain. After graduating from Juniata College in 1982, she volunteered to teach English in Comayagua, Honduras. In the spring of 1984, she spent a semester at the Université de Caen, Normandy, France, before entering the foreign language Masters program at West Virginia University. She graduated in 1986 with two years of teaching ESL, Spanish, Humanities and conducting academic advising for the College of Arts and Sciences. Later that same year, she moved to Austin, Texas, to begin graduate studies in anthropology. She worked as an assistant instructor in the Spanish and Portuguese department for one year. From 1987 to present she has worked on the campus of Central Texas College in ESL, Spanish, and anthropology. In August 2002 she entered the Foreign Language Education Program at the University of Texas at Austin. In the summer of 2008 created a virtual campus in Second Life for language learning and experiential anthropological projects. Her most important life's work has been raising a daughter who is currently an anthropology major at the University of Texas, Austin, preparing for her junior year abroad in Spain.

Permanent email: LimaVolle@gmail.com

This dissertation was typed by the author.