Language Learning & Technology
http://llt.msu.edu/vol9num3/payne/
September 2005, Volume 9, Number 3
pp. 35-54

SYNCHRONOUS CMC, WORKING MEMORY, AND L2 ORAL PROFICIENCY DEVELOPMENT

J. Scott Payne & Brenda M. Ross
The Pennsylvania State University

ABSTRACT

Recently a number of quasi-experimental studies have investigated the potential of a cross-modality transfer of second language competency between real-time, conversational exchange via text and speech (Abrams, 2003; Beauvious, 1998; Kost, 2004; Payne & Whitney, 2002). Payne and Whitney employed Levelt's (1989) model of language production and concepts from working memory as a rationale for a hypothesized connection between synchronous computer-mediated communication (SCMC) and second language (L2) speech and as a basis for predicting the differential contributions of SCMC to the L2 oral proficiency development.

This study extends the psycholinguistic framework reported in Payne and Whitney (2002) with discourse and corpus analytic techniques to explore how individual differences in working memory capacity may affect the frequency of repetition and other patterns of language use in chatroom discourse. Working memory capacity was measured by a reading span and nonword repetition test. Oral proficiency was measured with a speaking task that solicited a 5-minute speech sample and was scored based on a holistic scale. The data collected from 20 chat sessions were analyzed for occurrences of repetition and relexicalization, as well as language output measures. Findings suggest a connection between working memory and language output as measured in this study.

INTRODUCTION

Expressing oneself effectively and appropriately during oral conversational exchange with native or expert speakers of a target language represents for many learners and teachers the ultimate goal of language instruction. As a result, second language acquisition researchers have grappled with understanding, operationalizing, and testing oral proficiency (Hagen, 1990; Lantolf & Frawley, 1985; 1988; 1992; Kenyon & Malabonga, 2001; Liskin-Gasparro, 1984; Norris, 2001; Omaggio, 1983) while language teachers, either in accordance with national initiatives (i.e., American Council of Teachers of Foreign Languages [ACTFL] Proficiency Guidelines) or independently, have worked to develop pedagogical approaches and instructional activities that promote L2 speaking ability. The tacit assumption underlying these efforts has been that speaking is improved by practice speaking -- in a variety of situational contexts and on a range of topics with diverse socio-pragmatic requirements. However, emerging evidence from synchronous computer-mediated communication (SCMC) research suggests that real-time, conversational exchange via text may indirectly develop L2 speaking ability (Abrams, 2003; Beauvois, 1997; Payne & Whitney, 2002) and more importantly, may provide an extra benefit to learners with specific memory capacity limitations (Payne & Whitney). The present study builds upon this previous work with the goal of better understanding the characteristics of SCMC discourse and how they may facilitate a cross-modality transfer of language skill from SCMC to speech.

In the early 1990s, reports of the application of a new computer-mediated communication (CMC) technology enabling individuals to exchange text messages over a computer network in real-time, known as *chatting,* began to emerge in the second language research and teaching literature (Beauvois, 1992; Chun, 1994; Cononelos & Oliva, 1993; Kelm, 1992; Nicholas & Toporski, 1993). While none of these early accounts examined the transfer of skill from chatting to speaking directly, the strong resemblance

between the types of utterances generated in a chatroom and what would be said in a spoken conversation (Chun, p. 29), prompted researchers to speculate about a possible transfer of language skill from the real-time production of text to conversational L2 speech (Beauvois, 1992; Chun, 1994; Kelm, 1992). It was not until later, however, that the question of cross-modality transfer of language skill was addressed experimentally.

Researchers have compared SCMC and face-to-face discussion on a number of dimensions, including the investigation of the effectiveness of SCMC as a preparatory activity for face-to-face (F2F) discussion. These studies have been cross-sectional in nature, frequently comparing the quantity and nature of linguistic output during one chat session as compared with F2F discussion. Findings from these studies suggest that (a) students often produce more language in a chatroom than F2F settings (Abrams, 2003; Kern, 1995; Warschauer, 1996); (b) students tend to produce more complex language when chatting (Bölke, 2003; Kern, 1995; Warschauer, 1996) including more accurate usage of past-tense morphological markers (Salaberry, 2000); (c) there is greater equity in participation among students in a chatroom (Chun, 1994; Freiermuth, 1998, 2001; Kern, 1995; Warschauer, 1996); and (d) students exhibit improved attitudes towards foreign language learning as a result of chatting (Beauvois, 1992; Chun, 1994; Kern, 1995; Warschauer, 1996). Only a few studies to date have been more longitudinal in nature (one semester in length) and examined oral proficiency development as the result of numerous chat sessions during a semester. In the next section, we will review these investigations.

LONGITUDINAL STUDIES OF SCMC AND L2 ORAL PROFICIENCY

Beauvois' (1997) examined the difference in average oral exam scores of 83 participants randomly assigned to four sections of third-semester French (two sections in the experimental condition and two sections in the control condition). Based on *French in Action* (Capretz, Abetti, & Wylie, 1987) and a selection of short stories from *Les récrés du Petit Nicolas* (Sempé & Goscinny, 1997) all groups received 3 hours of instruction per week consisting of (a) a weekly quiz followed by a presentation of the next chapter's video teleplay (Day 1), (b) exploitation of the video material using the textbook's content and personal questions (Day 2), and (c) the pedagogical part of the video and workbook exercises (Day 3). The focus of the study was on student achievement in oral skills as result of the communicative activities associated with Day 2 in the sequence. Of the three instructors involved in the study, the first instructor taught one experimental and one control section, the second instructor taught one experimental section, and the third instructor taught one control section. The participants took three oral exams, one each in February, March, and May of 1995, and were evaluated by the instructors using a 20-point, holistic scale. Chat sessions for the experimental condition were conducted with the *Interchange,* the *Daedalus* LAN communication module. Results from a *t* test ($t = 2.20$, $p = .03$) suggested a greater improvement in oral skills for participants in the chat condition. While results from this study were favorable, there were also some limitations to the design. First, it is not possible to rule out a "teacher effect" or instrumentation threat to internal validity since only one of the three instructors taught both an experimental and control section. A second concern relates to the scoring procedure for the oral exams. Since the instructors evaluated their own students as opposed to employing independent raters, it was not possible, as a part of this design, to establish inter-rater reliability among raters.

An investigation by Abrams (2003) examined the effects of synchronous and asynchronous CMC and a control condition (no CMC) on student oral performance during three class discussions in a large-group discussion format. In this quasi-experimental design, performance was measured in terms of the number of words and communication units (c-units), lexical richness, lexical density, and the syntactic complexity of language generated during both synchronous and asynchronous CMC sessions and class discussions. Findings from this study indicate that synchronous CMC is a more effective preparatory activity for whole-class discussion than either asynchronous discussion or small-group and pair-work activities in a F2F classroom based on gain scores in the number of c-units ($F = 5.59$, $p = .005$) and words ($F = 10.02$, p

= .000) produced. For lexical richness, lexical density, and syntactic complexity no significant differences were found between the groups.

Kost (2004) conducted a quasi-experiment, where participants in the experimental condition engaged in 12 chat sessions during a 15-week semester. Oral and written proficiency were measured at the beginning and end of the semester. The oral proficiency test consisted of a 5-10 minute interview that was audio taped and scored with an analytic scale developed by the researcher. The oral posttest was counted as an oral exam. A timed (10 minutes), in-class writing sample was collected at the beginning and end of the semester as a measure of written proficiency with slightly different topics for the pre- and posttest. The treatment was a two-stage activity: Participants conducted a Web research activity followed by a role-play. In the control condition, the role-play was performed F2F in the classroom, whereas role-plays for participants in the experimental condition were performed in a chatroom. No significant differences were found between groups with respect to oral or written proficiency development.

It is important to note that in the context of such comparisons, finding no significant differences is not a "non-result" from a pedagogical perspective. Achieving equivalent development in oral skills with reduced F2F oral interaction should be considered a positive result. In light of the previously mentioned assumed superiority of F2F interaction for the development of oral proficiency, any comparison to other methods, at least conceptually, constitutes a directional hypothesis in favor of the face-to-face condition. Following this rationale, the finding of "no significant differences" could be posited as a rejection of the hypothesis that face-to-face is superior. Alternatively, such findings could be reformulated as "oral proficiency development was the same for the experimental condition with reduced face-to-face oral practice."

A study conducted by Payne and Whitney (2002) tested the hypothesis that the spontaneous production of text in a chatroom would develop the same cognitive mechanisms underlying L2 speech. The participants were 58 volunteers from four sections of fourth semester Spanish. In a pretest, posttest quasi-experimental design, participants in the experimental condition interacted in a chatroom two out of the 4 days of instruction per week over a 15-week semester for a total of 20 chat sessions. Participants in the control condition received the customary 4 days per week of F2F classroom instruction. Oral proficiency was measured using a 50-point holistic scale administered by independent raters with the pretest occurring during the third and fourth weeks of the semester (inter-rater reliability was .86) and the posttest during the final week (inter-rater reliability was .94). The curriculum was the same for all groups with participants engaging in the same activities on the same days across conditions. The only difference was that the activities for participants in the experimental condition took place in a chatroom 2 days per week. In order to statistically control for a possible "teacher effect," two instructors were involved in the study, each teaching one control and one experimental section. Findings from an ANCOVA with pretest as a covariate showed greater gains in oral proficiency for participants in the experimental condition (F = 3.96, p = .05).

Two important differences that exist between Payne & Whitney (2002) and other comparative studies of F2F and chatroom conversation are: (a) the use of Levelt's (1989) model of language production as a basis for making principled predictions about the cross-modality transfer of language skill between chatting and speaking and (b) the inclusion of concepts from working memory theory to account for individual differences in oral proficiency development. Since the psycholinguistics of language production are not the subject of the study reported in this paper, an extensive discussion of Levelt's and other models of language production is not warranted; however, a brief review of its potential for explaining the chat-to-speech connection may provide some useful background.

With Levelt's (1989) model of language production, utterances begin as pre-verbal thought in the *conceptualizer*. Once the communicative intention is generated, it emerges as a preverbal message and enters the *formulator* where lexical access and grammatical and phonological encoding occurs. After the

formulator is finished with the utterance-to-be, there are two options: either the message enters the *articulator* triggering speech-motor functions to produce the utterance, or it is monitored internally for accuracy and appropriateness by cycling back through the *speech comprehension system* as sub-vocalized internal speech. Based on Levelt's model, Payne and Whitney (2002) argued for the psycholinguistic near-equivalency of chatting and speaking with the only substantial cognitive difference being in the involvement of the speech-motor system (*articulator*).

The second difference was the application of working memory theory to second language SCMC research. Payne and Whitney (2002) found that the oral proficiency gains of the participants with lower phonological working memory capacity (PWMC), as measured by nonword repetition, boosted the mean for the experimental group to the point that it exceeded the control group. Based on this finding, the authors concluded that the chatroom had provided low-span individuals with a means of compensating for limitations in working memory and assisted them in staying engaged in the conversation, which had resulted in increased practice and the ultimate transfer of language skill. In recent years, SLA researchers have become increasing interested in studying the effects of individual differences in working memory capacity on second language learning processes. In the next section, we review findings from this growing body of research.

WORKING MEMORY AND SECOND LANGUAGE LEARNING

Basic language processes such as lexical access and syntactic processing require conscious attention, especially for beginning and intermediate foreign language learners. While attending to these processes, activated information needed for further processing must be maintained in memory. The system or mechanism supporting the temporary maintenance of information during the performance of complex cognitive tasks, such as second language comprehension and production, is known as working memory (Baddeley & Hitch, 1974; Daneman & Carpenter, 1980). Two of the most frequent measures of working memory are the reading span, originally developed by Daneman and Carpenter, and the nonword repetition test. The reading span is a measure of an individual's ability to process and temporarily store information in memory and has provided the empirical underpinnings of Just and Carpenter's (1992) constrained capacity theory of working memory. This theory posits a unitary pool of attentional resources that are allocated to both processing and storage. This theory differs from Baddeley's (1986, 2000) theory that depicts working memory as a multi-component system. In Baddeley's model there are four components: *phonological loop, visuo-spatial sketchpad, central executive,* and *episodic buffer.* Most relevant to the present study are the phonological loop and the central executive. The phonological loop (PL) stores sound information in memory and maintains these memory traces via sub-vocal rehearsal. The PL is a slave system of the central executive, the attentional control system. Some researchers consider the reading span as a measure of central executive capacity (Engle, Kane, & Tuholski, 1999; Baddeley, 2003). The nonword repetition task is the most popular measure of the phonological loop or phonological working memory.[1]

In studies of first language, working memory capacity has been shown to account for individual differences in text comprehension (Daneman & Carpenter, 1980) and the ability to resolve syntactic ambiguity (Just & Carpenter, 1992; MacDonald, Just, & Carpenter, 1992). In both of these instances, individuals with higher memory span are superior at comprehending text and are able to maintain multiple interpretations of a text, resulting in better ambiguity resolution. Processing syntactically complex sentences (object-relative vs. subject-relative) is also more difficult for low-span readers and negatively impacts their comprehension (King & Just, 1991). Working memory capacity has also been correlated with verbal fluency (Daneman, 1991) and the ability of individuals to utilize contextual clues in text for learning novel words (Daneman & Green, 1986).

In studies of second language learning, Harrington and Sawyer (1992) and Geva and Ryan (1993) found that L2 reading skill is highly correlated with L2 working memory span. Abu-Rabia (2003) found

significant correlations between working memory span and L2 writing proficiency as measured by the *Test of Written Language.* Fortkamp (1999) found significant correlations between the speaking and reading span tasks and measures of L2 fluency.

Research on the role of the phonological loop in language learning indicates that PWMC is related to foreign language vocabulary learning in laboratory (Papagno, Valentine, & Baddelely, 1991; Papagno & Vallar, 1995) and classroom settings (Cheung, 1996; Service, 1992; Service & Kohonen, 1995) with high-span learners acquiring foreign-sounding words more easily than learners with more limited capacity. Ellis (1996) has argued that acquiring "chunks" forms the basis of language acquisition and that PWMC is instrumental in this process. Ellis and Schmidt (1997) found not only that short-term memory for utterances predicts later long-term memory for the same utterances; they also demonstrated that short-term memory representations of utterances is important for learning syntax. In a further study examining the role of repetition in vocabulary learning, Ellis and Sinclair (1996) required native-speakers of English learning Welsh as a second language to produce the Welsh equivalents of English words after a learning phase, where subjects had learned the English translations of Welsh utterances. There were three treatment conditions in this experiment. In the first experimental condition, subjects were instructed to repeat aloud the Welsh utterance after hearing it. In condition two, subjects were instructed to learn the Welsh items under articulatory suppression (subjects were required to repeat a number sequence aloud with the goal of blocking access to the PL). In the control condition, subjects were not given any specific instructions. Results demonstrated superior performance by subjects in the repetition condition, followed by the control condition and finally the articulatory suppression condition.

Based on the work of Ellis and colleagues showing the importance of repetition and PWMC in vocabulary acquisition, a productive next step would be to examine how repetition in conversational contexts impacts language development. Due to similarities between L2 chatroom discourse and conversational speech, chat transcripts may provide a valuable source of data for this inquiry. One phenomenon of repetition in conversational language use has received empirical attention from a number of theoretical and methodological orientations. The next section reviews some of this work.

REPETITION IN CONVERSATIONAL LANGUAGE USE

The coordinated act of conversation requires interlocutors to balance and negotiate topic and lexical knowledge together with the flow of discourse to produce a successful communicative interaction. A common characteristic of conversational discourse is the reuse of linguistic materials from prior talk in current talk (Schenkein, 1980). This act of recycling one's own language or the language of conversational partners has been the object of study among sociolinguists, psycholinguists, and corpus linguists.

Levelt and Kelter (1982) conducted a series of six experiments to examine the influence of a question's surface form on the format of an interlocutor's response from a psycholinguistic perspective. The notion underlying this investigation was that "previous talk sets up a more or less abstract frame in the mind of an interlocutor, which is then used in the formulation of the next turn" (p. 79). In the first experiment, they demonstrated that when interlocutors are presented with a question phrase in either of two equivalent forms, based on the absence or presence of a preposition in Dutch, they most frequently choose to formulate their responses using the same format as the question. The authors refer to this phenomenon as the "correspondence effect." Experiments 2-5 manipulated the question formats to include distracting information in the form of multiple questions under more everyday circumstances by employing Clark's (1979) telephone technique to determine whether memory was involved in this effect. Findings indicated that relevant information pertaining to the format of the question was maintained in working memory during the preparation of the answer. While question formats were controlled for "naturalness," they were not controlled for frequency of use. Furthermore, working memory span was not measured and included as an independent variable.

Accounts of repetition and a related linguistic phenomenon, relexicalization, have been documented in corpus linguistic studies. McCarthy (1998) argues that repetition often occurs in discourse when interlocutors attempt to agree on a topic. According to McCarthy, exact repetition often suggests a desire to deliberately stall the topic progression and can be interpreted as a refusal to communicatively accommodate. However, when repetition is accompanied by relexicalization (the re-casting of content with near-synonymous words), convergence is achieved.

Both of these studies examined repetition in the conversational speech of native or expert speakers. The extent to which these documented linguistic phenomena may manifest themselves in the conversational speech of L2 learners as well as serving other purposes is unknown. Specifically, do repetition and relexicalization play a role in interlanguage development and does their use relate to conversational and topical context or possibly even the working memory profile of the learners?

The research reviewed above posits the importance of working memory in second language vocabulary learning; the acquisition of syntax, text comprehension, and production; and the development of speaking skills. Other studies have examined the role of repetition and relexicalization in first language (L1) conversational language use. To date, no study has examined repetition in L2 chatroom discourse in an effort to shed light on the communicative functions of repetition in L2 learner language. Of further interest is whether individual differences in working memory capacity may affect the frequency of repetition and other patterns of language use. The purpose of the present study is to examine these linguistic data for the lexical, morphosyntactic, and pragmatic choices interlocutors made when co-constructing conversational discourse in the chatroom with special attention to the role that individual differences in working memory capacity may play in explaining patterns of language use. Based on previous psycholinguistic and corpus linguistic research, analyses will focus on the repetition of single lexical items and multi-word sequences by interlocutors and the use of relexicalization, as well as language output measures.

RESEARCH PURPOSE

In Payne and Whitney (2002), the chatroom was employed as a training mechanism to promote the automatizing of the cognitive processes underlying spontaneous L2 language production (e.g., lexical access, grammatical and phonological encoding) with the hypothesis that this practice, and resultant development, would transfer across modalities to speaking. Since the language generated in the chat sessions was not the subject of analysis in the first study, the focus of the present investigation is to examine the patterns of language use as evidenced in the chat transcripts (experimental group only) with the goal of better understanding the interplay between individual differences in working memory capacity, SCMC, and cross-modality transfer of skill from chatting to oral proficiency development.

RESEARCH QUESTIONS

1. Does the frequency of repetition and relexicalization in L2 SCMC change over the course of the semester?
2. Do individual differences in working memory capacity (phonological and executive function) play a role in the frequency of repetition and relexicalization in L2 SCMC?
3. Do individual differences in working memory capacity (phonological and executive function) modulate language output as measured by the average number or words, utterances, and turns generated per chat session?
4. What is the relationship between individual differences in working memory and oral proficiency development?

METHODS

Participants

The participants consisted of 24 volunteers (2 males and 22 females ranging in age from 18-26) from two sections of third-semester Spanish at a medium-sized university in the western United States. Students received extra-credit points equivalent to one-third of a letter grade for participating in the study. Those students who did not wish to participate in the study had the option of writing a one thousand-word essay on the cultural topic of their choice to receive the extra-credit points.

Instructional Treatment

Both sections of third-semester Spanish received 4 hours of hybrid or blended instruction per week (2 days online in the chatroom, 2 days in the classroom). Instruction had a communicative orientation, based on the textbook *Nuevos Destinos* (Medina, 1997) and its accompanying video series. The curriculum was uniform across both sections with students engaging in the same activities on the same days. The chatroom activities included small-group discussion, role-plays, discussions of assigned texts and video content, as well as information gap activities. Each class engaged in 20 different 50-minute chat sessions over the course of the 15-week semester with 3-4 separate chatrooms per course for a total of 150 chatroom transcripts.

Most discussions were based on questions given to the students by their instructors (both sections were given the same set of questions to discuss). These questions mostly elicited language taken out of texts, especially during the first third of the chat data sets. However, with time, students seemed to become more confident and began moving more and more away from their texts. The instructors also used jigsaw reading activities as triggers for discussion on two occasions. During these activities students rarely strayed from the language of the readings they were dealing with. These two activities took place near the end of the first half of the semester. Topics covered were as follows:

First half of the semester
- Daily routines
- Places of interest
- Occupations
- Family relationships in American and Hispanic families
- Bilingualism and bilingual education
- Obesity and food
- Childhood memories

Second half of the semester
- Means of transportation
- Gender differences
- Shopping
- Plans for the future
- Legends
- Vacations

Both halves of the semester
- Discussion of a didactic videotape that students viewed every week
- Diverse grammatical points
- Culture and traditions of Hispanic countries and of the US

Overall, the topics selected by the instructors were all conducive to good discussion. However, two topics seemed to be especially popular: talking about childhood memories (end of first half of the semester) and talking about plans for the future (end of second half of the semester). It must be noted, however, that both sections suffered technical difficulties during the beginning and end of the semester, which may have affected the quantity of production during chat times.

Materials and Procedures

Oral proficiency was measured with a task requiring participants to select one of four envelopes containing instructions for a speaking task written in English. Participants were required to read the instructions and then speak in Spanish for approximately 5 minutes. If participants ran out of things to say on a particular topic, they selected a new task and began again. The objective was to obtain a 5-minute speech sample. The role of the examiner was to listen but not to interview the speaker. Two examiners

(one native speaker and one non-native speaker; both were female) administered the speaking tests. The examiners were told to think of someone they know who is a very fluent non-native speaker of Spanish and consider that individual's language ability as a perfect score (the 50 point maximum).

Two working memory measures were included in this study: (a) nonword repetition task, as a measure of phonological working memory and (b) reading span, a measure of executive function or what Baddelely (1986) refers to as the central executive. The nonword repetition task developed for this study differed from conventional versions in that it was recognition-based and administered online. In this test, participants listened to an audio file of eight pseudo-words read with a one-second interval between words. After listening to the audio file, participants clicked on a button to view a screen containing 16 pseudo-words, 8 of which were articulated in the audio file. Students selected the eight words they believed to have heard by clicking on the checkbox next to each word. The participants could take as much time as they needed to make their eight selections. After clicking the submit button, the next audio clip would load, ready to be played. The nonword repetition task consisted of three sets of eight pseudo-words with a total possible score of 24.

The second working memory test administered was a reading span test. This measures an individual's ability to simultaneously process and store information from one's immediate environment in memory. In the recognition- and Web-based version of the reading span administered in this study, participants viewed 15 sets of sentences with a total of 60 sentences. For each set, one sentence was displayed after another in 7-second intervals until all of the sentences in a set had been viewed. While reading each of the sentences, participants had two tasks: (a) they had to decide whether the sentence made sense or not by selecting the radio button corresponding to their estimate of the sentence's sensibility, and (b) they had to remember the final word of each sentence. After all sentences in the set had been viewed, participants clicked on a button to receive a screen of words with adjacent checkboxes that contained both targets and distractors. For each to-be-remembered word there were two distractors (e.g., for sets containing five sentences there was a total of 15 words). Distractors were of two types: (a) same semantic category (e.g., if the target word was "girl", the distractor could be "woman"), or (b) final words from sentences in previous sets. Participants simply selected the words they identified as being final words by clicking the checkbox next to the word. One point was awarded for each final word recalled for which a correct sensibility judgment was made. The reading span score was a raw score with 60 points possible. All working memory tests were recognition- and Web-based with a database back-end, enabling automatic scoring and calculating of the results.

The reading span and nonword repetition tests were administered in a computer laboratory during the second week of the 15-week semester. The speaking pretest was administered during week three with the posttest occurring during the fifteenth week of the study. The instructional treatment that is the focus of this study consisted of 20 chat sessions where participants engaged in a variety of structured and open-ended tasks ranging from personal themes, role-plays, and discussions of textbook readings to the negotiation and co-construction of understanding of video content. Instructors participated in the chat discussions, posing questions to reinvigorate dialog when necessary and occasionally providing corrective feedback.

Data Analysis

A median split was performed to divide participants into high- and low-span groups for both the reading span and the nonword repetition tests, making working memory span an independent variable. Individuals whose score fell on the median were assigned to either the high or low group depending on which side of the mean their score fell. This resulted fortuitously in an even split with 12 participants in each of the span groups for both working memory measures. The changes in oral proficiency based on working memory span group for both working memory measures were calculated with an ANCOVA using the oral pretest score as a covariate to factor out pre-existing proficiency differences from the posttest score.

The chatroom data were divided into two periods representing the first and second halves of the semester in order to measure potential developmental changes in language use. Frequencies were tabulated for average number of words, utterances, and turns generated per chat session per participant, as well as repetition and relexicalization (all dependent variables are defined later). Since only one of the researchers is knowledgeable in Spanish, this researcher carried out all coding of data. Reliability in coding procedures was established by discussing representative candidate instances of each code to control for threats to interpretive validity. The Wilcox Signed Ranks and Mann-Whitney *U* tests, the non-parametric equivalents to the *t* test, were calculated to compare groups based on dependent measures for language output, repetition, and relexicalization, with span for the individual working memory measures as the independent variable. The Kruskal-Wallis, the non-parametric equivalent to the one-way ANOVA, was calculated to compare groups for the secondary analysis of interaction between working memory measures. Since non-parametric tests for group comparisons do not permit the use of covariates, gain scores formed the basis of the developmental comparison between the first and second halves of the semester. The alpha level for all group comparisons was set at .05.

Definition of Terms

The chat data was analyzed mainly with the intention of looking at the role of repetition and relexicalization in the production of language during interactions. Other artifacts of language use, including corrections and the use of "Spanglish" and of invented words, were prominent in the chatting of some students, but are not analyzed for the present study.

For this paper we have defined repetition as the repetition of a single word or phrase from an earlier turn by a peer or by the instructor. For example, one student may have brought up a new word, which was subsequently repeated in a later turn by another student, as in Example 1.

Example 1

A: *¿Tienes muchas **admiradores** Chad?*
Do you have many admirers Chad?
…
B: *Si, especialmente **admiradores** que estan chicas bonitas.*
Yes, especially admirers that are pretty girls.

In Example 2, student B repeated the phrase *"el programa es muy mal"* (the program is very bad) with inversion of the subject and verb for question formation.

Example 2

A: *Sí hay una programa de deucación bilingüe donde yo vivo pero la programa es muy mal.*
Yes there is a bilingual program where I live, but it is very bad.
B: *¿Por qué es el programa muy mal?*
Why is the program very bad?

Relexicalization is similar to repetition in that a similar *idea* is repeated, but the structure and/or one or more of the main grammatical words or phrases (nouns or verbs) are different. For example, in Example 3, the students were both talking about the weather in Chile, but the subject of the second sentence is no longer "the weather," but Chile; the structure of the sentence has been changed.

Example 3

A: *Yo quiero ir a Chili ahora porgue **el tiempo de Chili es calor.***
I want to go to Chile because the weather in Chile is hot.
B: *Es verdad que **en Chile hace calor**, pero en las montañas hace mucho frío.*
It is true that in Chile it is hot, but in the mountains it is cold.

Example 4

A: ***Tienen que*** *crear un diálogo entre las personas de una familia latina.*
You have to create a dialogue between the members of a Latin family.
B: ***Necisitamos*** *crear un diálogo de una familia latina y estadounidense.*
We need to create a dialogue of one Latin family and an American [one].

Example 4 is a case of a sentence that is both a repetition of an earlier turn and a relexicalization of it. Student B has repeated the instructor's (in this case, "A") instructions, but has changed the main verb from "have to" to "need," thus effectively keeping the main idea, but changing the structure enough for the sentence to be novel.

RESULTS

Changes in the Frequency of Repetition and Relexicalization

To address research question one, first the descriptive statistics were calculated for the ratio of total repetitions and relexicalizations by total chat turns occurring in the first and second halves of the semester (see Table 1). Next the repetition and relexicalization ratios for the first and second half of the semester were compared using the Wilcoxon Signed Ranks Test (see Table 2).

Table 1. Descriptive Statistics for Repetition and Relexicalization

	N	Mean	*SD*
Repetition - 1st half	24	.214	.123
Repetition - 2nd half	24	.084	.052
Relexicalization - 1st half	24	.045	.040
Relexicalization - 2nd half	24	.024	.019

Table 2. Wilcoxon Signed Ranks for Repetition and Relexicalization

		N	Mean Rank	Sum of Ranks
Rep. 2nd half -- Rep. 1st half	Negative Ranks	24	12.50	300.00
	Positive Ranks	0	0.00	0.00
	Ties	0		
	Total	24		
Rel. 2nd half -- Rel. 1st half	Negative Ranks	18	13.06	235.00
	Positive Ranks	5	8.20	41.00
	Ties	1		
	Total	24		

Results from the Wilcoxon Signed Ranks test indicate that there was a significant difference between the first and second halves of the semester in terms of the frequency of repetition (Z = -4.286, Asymp. Sig. = .000) and relexicalization (Z = -2.950, Asymp. Sig. = .003) with participants more frequently recycling single- and multi-word elements of their interlocutors' language in their own chat discourse, as well as in the form of phrasal transformation known as relexicalization, during the first half of the semester.

Working Memory and the Frequency of Repetition and Relexicalization

As an initial step in addressing research question 2, descriptive statistics were generated for the gain scores for repetition and relexicalization from period 1 to period 2 (period 1 mean frequency/chat turn ratio was subtracted from the period 2 frequency/turn ratio) with a breakdown based on low- and high-span groups for both the reading span and nonword repetition tests (see Table 3). The comparison of mean frequencies from the first half to the second half of the semester shows a consistent decline in the use of repetition and relexicalization for all groups.

Table 3. Descriptive Statistics for Frequencies of Repetition and Relexicaliation

		Reading Span		Nonword Repetition	
		Low	**High**	**Low**	**High**
Repetition	Mean	-.145	-.115	-.102	-.158
	SD	.117	.057	.046	.117
	N	12	12	12	12
Relexicalization	Mean	-.021	-.020	-.012	-.030
	SD	.055	.017	.025	.050
	N	12	12	12	12

Mann-Whitney *U* tests for each of the working memory measures were calculated to determine if there were differences between the groups based on span. No significant differences between high- and low-spans were found for both the reading span and nonword repetition with respect to changes in the frequency of repetition or relexicalization employed in the chatroom from the first to second half of the semester. These results suggest that the frequency of repetition and relexicalization in L2 chatroom discourse cannot be explained by working memory capacity limitations.

Working Memory and Language Output Measures

The descriptive statistics were calculated for the differences in frequencies between the first and second halves of the semester for the average number of words, utterances, and turns generated per chat session based on span group for the reading span and nonword repetition tests (Table 4).

Table 4. Descriptive Statistics for Gain Score Frequency Ratios on Language Output Measures

		Reading Span		Nonword Repetition	
		Low	**High**	**Low**	**High**
Words	Mean	38.45	104.30	104.20	38.56
	SD	78.23	64.24	64.53	78.09
	N	12	12	12	12
Utterances	Mean	6.15	16.15	13.77	8.52
	SD	7.31	11.53	12.79	7.88
	N	12	12	12	12
Turns	Mean	2.82	9.41	7.33	4.50
	SD	6.22	5.29	4.88	7.80
	N	12	12	12	12

Mann-Whitney *U* tests were calculated to compare the change in mean frequency per chat session of the language output measures based on working memory capacity (see Table 5). For phonological working memory (nonword repetition) significant differences were found between low- and high-spans (Z = -2.309, Asymp. Sig. = .021) with lower phonological working memory capacity students producing a greater number of words on average per chat session than high-span students. No significant differences were found between phonological working memory groups for mean utterances or turns generated per chat session. This finding suggests that students with low phonological working memory span exhibited a distinct chat style with longer utterances (sentences), but not more of them on average per turn. The implications of this behavior will be pursued further in the discussion section.

For the comparison of reading span groups, the findings reveal significant differences between low- and high-span participants on the average number of words generated per chat (Z = -2.194, Asymp. Sig. = .028), as well as the average number of utterances (Z = -2.714, Asymp. Sig. = .007) and turns (Z = -2.309, Asymp. Sig. = .021) contributed to a chat session.

Table 5. Mann-Whitney *U* Test for Language Output Measures

	Avg. Words/Chat		Avg. Utterances/Chat		Avg. Turns/Chat	
	Reading Span	**Nonword Repetition**	**Reading Span**	**Nonword Repetition**	**Reading Span**	**Nonword Repetition**
Mann-Whitney *U*	34	32	25	52	32	52
Z	-2.194	-2.309	-2.714	-1.155	-2.309	-1.155
Asymp. Sig.	.028	.021	.007	.248	.021	.248

Based on the descriptive statistics in Table 4, there appears to be an inverse relationship between capacity and working memory measure with high-spans on the reading span generating more language than the low-spans and results for the nonword repetition test depicting the reverse scenario. Since all participants took both working memory tests and not all participants scored either low or high on both measures, a secondary analysis was undertaken to determine if there was an interaction between working memory test and span level. This was accomplished by dividing the 24 participants into four groups: (a) nonword repetition-low and reading span-low, (b) nonword repetition-low and reading span-high, (c) nonword repetition-high and reading span-low, and d) nonword repetition-high and reading span-high. The descriptive statistics for language output gain score frequencies are reported in Table 6.

Table 6. Descriptive Statistics for Language Output Gain Score Frequencies

		Low-Low	**Low-high**	**High-low**	**High-high**
Words	Mean	85	123.39	-8.10	85.22
	SD	79.52	44.22	44.06	79.05
	N	6	6	6	6
Utterances	Mean	7.81	19.73	4.48	12.57
	SD	8.13	14.45	6.69	7.26
	N	6	6	6	6
Turns	Mean	4.94	10.53	.71	8.29
	SD	4.84	3.19	7.04	6.96
	N	6	6	6	6

The groups were compared using the Kruskal-Wallis test with alpha level set at .05 (see Tables 7 and 8). Results indicate a significant difference between the groups for the mean number of words and utterances generated from the first half to the second half of the semester and a non-significant trend for frequency of turns. Pair-wise post-hoc comparisons found group three (nonword repetition high -- reading span low) to be significantly lower than the other groups in average total words generated per chat session. The pair-wise comparisons based on average utterances produced per chat session found differences between groups two and three, but no other differences between groups emerged.

Table 7. Ranks for Language Output Gain Scores by Working Memory Span Group

	Span Group	*N*	Mean Rank
Words	1	6	13.50
	2	6	18.17
	3	6	5.17
	4	6	13.17
	Total	24	
Utterances	1	6	10.33
	2	6	18.00
	3	6	6.83
	4	6	14.83
	Total	24	
Turns	1	6	11.17
	2	6	17.17
	3	6	7.17
	4	6	14.50
	Total	24	

Group 1 = low nonword rep. - low reading span
Group 2 = low nonword rep. - high reading span
Group 3 = high nonword rep. - low reading span
Group 4 = high nonword rep. - high reading span

Table 8. Kruskal-Wallis Test for Language Output Gain Score Frequencies

	Words	Utterances	Turns
Chi-Square	10.480	8.700	6.720
df	3	3	3
Asymp. Sig.	.015	.034	.081

Working Memory and Language Proficiency Development

The fourth research question examines a hypothesized relationship between working memory capacity and the development of L2 speaking ability. In the previous study (Payne & Whitney, 2002) the comparison was made between a control condition and the experimental chatroom condition. The present study seeks to investigate this hypothesized relationship further in an effort to determine if individual differences in working memory capacity may further explicate development in oral proficiency among students interacting in a chatroom for half of their instruction time. To address this question, two ANCOVAs were calculated: one with phonological working memory as the independent variable with two levels (high- and low-span) and the other with executive function (measured by reading span) also with two levels (high- and low-span). The dependent variable for both statistical procedures was the oral proficiency posttest score with the pretest score as a covariate with the alpha level set at .05.

Table 9. Descriptive Statistics for Oral Proficiency Scores by WM Test

Group	Pretest Mean	Pretest *SD*	*N*	Posttest Mean	Posttest SD
Nonword - low	20.63	6.71	12	30.50	5.08
Nonword - high	20.79	7.11	12	34.79	6.33
Reading Span - low	19.75	5.60	12	30.71	5.89
Reading Span - high	21.67	7.88	12	34.58	5.74

For nonword repetition, the results from the ANCOVA suggest that the oral proficiency of students with greater PWMC improved more over the course of the semester than for their low-span counterparts (F = 4.325, p = .050). Results for the comparison based on reading span indicate no significant differences between the groups with respect to oral proficiency development (F = 2.059, p = .166).

To control for possible interaction effects between the two working memory measures (some individuals where high-spans in one measure and low-spans in the other), a secondary analysis was conducted using the same procedure as in research question three. The ANCOVA was calculated with working memory span group as the independent variable with four levels, oral proficiency posttest score as the dependent measure, and the oral pretest score as the covariate with the alpha level set at .05. Results indicate that there is no significant difference between the four working memory span groups in oral proficiency development.

DISCUSSION

The present study is an initial investigation into the relationships between individual differences in working memory capacity, patterns in L2 chatroom language use, and oral proficiency development. Several results emerged from this study that warrant further discussion. First, the use of repetition and relexicalization by third-semester Spanish learners as a strategy to facilitate communication in L2 chatroom discourse declined in frequency over time, an effect that was unrelated to working memory capacity. One plausible interpretation for this trend is that the frequency of repetition and relexicalization is inversely related to proficiency. In other words, as learners expand their lexicon and become more proficient in producing language spontaneously as required for conversational exchange, they reduce their reliance on repetition as a strategy for maintaining continuity in discourse. Pretest-posttest changes in oral proficiency and repetition/relexicalization frequencies were correlated in an effort to test this deduction (Kendall's τ = -.166 for repetition and -.044 for relexicalization). The weak relationship between gains in oral ability and frequency of repetition/relexicalization does not lend support to the notion of a proficiency-repetition connection in this study.

Another possibility is that these phenomena of language use in L2 chat are topic-related. During the first half of the semester, a number of the chatroom discussions were based on texts introducing topics for which students may not have had much background knowledge (e.g., bilingualism and bilingual education, obesity and food, relationships in American and Hispanic families). A lack of background knowledge coupled with a more focused discussion topic could have resulted in greater repetition and relexicalization as students endeavored to move the discussion forward and orient their comments to the assigned text and discussion questions. These conclusions suggest that L2 learners may use repetition for very different reasons than has been documented in L1 corpus linguistic research (e.g., McCarthy, 1998). Contrary to L1 evidence suggesting that repetition is employed in discourse to stall topic development, the L2 learners in this study may have used repetition to promote continuity and focus the discussion. In the second half of the semester, topics were more familiar and discussion questions were less guided. As a result, students generated more language and were possibly less inclined to employ repetition to encourage discussion.

A second finding of interest is the apparent difference in the chatting style of the low phonological working memory students. The low-span chat style was characterized by a greater number of words per utterance on average than was exhibited by high-span students. While L2 chat discourse has been characterized by longer, more elaborate constructions in comparison to F2F conversation (e.g., Kern, 1995; Warschauer, 1996), the use of longer, multi-sentence comments can also be an indicator that users are inexperienced in the SCMC communicative modality. For the low-span (PWMC) students participating in this study, this is not likely the case, due to the persistence of this behavior throughout the semester. Another interpretation consistent with the posited conversational "bootstrapping effect" of the chatroom, is that the low-span students were taking advantage of the reduced cognitive burden introduced

by the chatroom to produce more extensive and elaborate constructions; something they may have found difficult in a F2F setting. A more fine-grained syntactic analysis of the chat comments would be required to verify this interpretation.

A third finding comes from the secondary analysis of language output measures in research question two, examining a potential interaction between working memory measures. The results from this analysis revealed an interaction between executive function and phonological working memory that makes the interpretation of the role of working memory in L2 SCMC potentially more complex. The language output frequencies for Group 2 (low nonword repetition and high reading span) and Group 3 (high nonword repetition and low reading span) are of greatest interest since they represent a crossover in levels between the two tests. Based on the means for the average number of words, utterances, and turns per chat session, Group 2 exceeded all other groups; however this group was statistically different only from Group 3. One could speculate that Group 2 exemplifies the "bootstrapping effect" posited by Payne and Whitney (2002). Group 3, with the opposite working memory profile of Group 2, generated the least amount of language output of all four groups. It is plausible that these higher PWMC individuals did not benefit as much from the reduced cognitive processing burden of the chatroom and preferred the classroom environment for conversational activities. In any case, these findings suggest a potentially interesting interaction between the reading span and nonword repetition as measures of working memory capacity and their impact on performance in non-laboratory settings. Since no research to date has examined the combined effect, this may be a productive line of future research.

An additional candidate for further discussion is the role of individual differences in working memory in L2 oral proficiency gains. While all groups, based on the median split for each working memory measure, realized gains in their oral proficiency over the semester, significant differences were found only for phonological working memory with the high-span group out-performing the low-span group. This result is not surprising in light of the hypothesized role of PWMC in SLA. When considered together with the finding that low-span students produced considerably more language in the chatroom, the relationship between working memory and oral proficiency development lends further credence to the idea that low-span learners are using the chatroom as a compensatory mechanism.

LIMITATIONS

There are a couple of limitations in this study that should be mentioned. First, the ability to find differences between groups (power) was reduced by the small sample size. It is likely that the analyses examining the interaction between working memory measures especially suffered from this limitation. A further complication introduced by smaller samples is the overbearing effect that outliers can have on group mean scores. This can be a particular challenge when data are collected in a more naturalistic setting without the controls of a laboratory. The measures of language output, repetition, and relexicalization consisted of frequency counts in naturally-occurring L2 chatroom language use. Due to the nature of this data, a wide range of variability was evident by the large standard deviations reported in the descriptive statistics. This effect can work two ways: uncharacteristically low scores can mask potential effects whereas exceeding high scores can result in Type 1 errors (finding effects that do not exist). A larger sample size would increase power and ameliorate the effects of outliers on group means.

CONCLUSION

This study represents an initial attempt at using working memory as a lens for interpreting patterns of language use as learners co-construct discourse in a chatroom and their relation to L2 oral proficiency development. The language patterns that were the focus of this inquiry included repetition and relexicalization, as well as the average number of words, utterances, and turns per chat session. While individual differences in working memory could not be found to account for the drop in the frequency of repetition and relexicalization from the first to second half of the semester, a role for working memory

capacity was supported in language output. The findings of this study lend further support to the notion that the chatroom may provide a unique form of support to certain types of learners in developing L2 oral proficiency. Evidence for this "bootstrapping effect" was found in greater language output measures of students with lower PWMC. While these results suggest a relationship between working memory and chatroom language use, more extensive research with larger samples is needed. As a final observation it is important to consider these findings evidence that language learners will use the tools at their disposal to maximize their learning outcomes. Historically, computer-assisted language learning (CALL) as a field has focused on how software applications can improve vocabulary knowledge (explicitly and implicitly), listening and reading comprehension ability, and writing skills; analyze and correct pronunciation; and provide speaking practice. While the present investigation has addressed the question of learning outcomes, of equal importance is the proposal that technology-mediated language learning transforms the cognitive processing constraints of tasks in a manner that has direct benefit to certain types of learners (see Robinson, 2001, for further discussion). Continued research in this area stands to make a substantial contribution to our understanding of SLA processes as well as to second and foreign language pedagogy.

FUTURE RESEARCH

With respect to the L2 SCMC corpus data analyzed for this study, a number of subsequent analyses of other patterns in L2 SCMC language use (e.g., correction, Spanglish, invented words) and their relationship to individual differences in working memory capacity are in order. Further research is clearly needed to better understand the role of repetiton and relexicalization in L2 conversational language use in F2F and SCMC settings. As previously mentioned, performed syntactic analyses of the chat contributions based on phonological working memory span could shed additional light on the distinct "chat style" of low phonological working memory foreign language learners.

An emerging area of SLA research is the role that formulaic sequences play in interlanguage development. Wray and Perkins (2000) and Wray (2002) proposed a conceptual model, based on Locke (1993, 1995, 1997), for how formulae may function in L2 acquisition. In this model, early stages of acquisition are characterized by the learning of unanalyzed, formulaic sequences also known as "chunks." At the second phase of the model, once learners have amassed a database of chunks and as the stochastic patterns of those chunks in use become salient, learners begin to analyze and identify the constituent structure of these formulae. This results in the emergence of the grammatical system. As the focus shifts to analyzing chunks and learning to substitute constituents to create novel utterances, the acquisition of formulaic sequences drops. Exploring the use of formulaic sequences based on this model may be a very fruitful avenue for SCMC research with the added dimension of comparing chatroom and F2F conversational discourse for the appropriation and use of formulae.

A final promising possibility in SCMC research would be the use of time-series statistics for measuring change over time with the hope that change in frequency reflects interlanguage development. This technique would offer the benefits of numerous observation points and the ability to compare learner trajectories as opposed to mean frequencies based on arbitrary divisions of the data into time periods. This approach to analysis may be particularly well suited to SLA research where development is often characterized by U- or Omega-shaped curves instead of the linear relationships at the core of parametric statistics (see Norris & Ortega, 2003, for further discussion).

NOTE

1. In this article, we use phonological working memory capacity and phonological short-term memory synonymously.

ABOUT THE AUTHORS

J. Scott Payne (PhD, Washington State University) is Assistant Director for Technology and Research at the Center for Language Acquisition and Co-Director of the Technology Project under CALPER at The Pennsylvania State University. His research focuses on the role of individual differences in working memory and attention in CALL.

E-mail: jspayne@psu.edu

Brenda M. Ross is a doctoral student (ABD) at the Pennsylvania State University. She is currently a Senior Lecturer in TESL at the University of Texas at Tyler. Her research interests include discourse analysis, narratives, and identity, and the use of corpora for ESL instruction.

E-mail: bmr157@yahoo.com

REFERENCES

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

Abu-Rabia, S. (2003). The influence of working memory on reading and creative writing processes in a second language. *Educational Psychology, 23*(2), 209-222.

Baddeley, A. D. (1986). *Working memory.* Oxford, England: Clarendon Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*(11), 417-423.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36,* 189-208.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In S. Dornic (Ed.), *Recent advances in learning and motivation* (Vol. VIII, pp. 47-89). New York: Academic Press.

Beauvois, M. (1992). Computer assisted classroom discussion in the foreign language classroom: Conversation in slow motion. *Foreign Language Annals, 25*(5), 455-464.

Beauvois, M. H. (1997). Write to speak: The effects of electronic communication on the oral achievement of fourth semester French students. In J. Muyskens (Ed.), *New ways of learning and teaching: Issues in language program direction* (pp. 93-115). Boston: Heinle & Heinle.

Beauvois, M. H. (1998). Conversations in slow motion: Computer-mediated communication in the foreign language classroom. *Canadian Modern Language Review, 54*(2), 198-217.

Böhlke, O. (2003). A comparison of student participation levels by group size and language stages during chatroom and face-to-face discussions in German. *CALICO Journal, 21*(1), 67-87.

Capretz, P., Abetti, B. & Wylie, L. (1987). *French in action: A beginning course in language and culture.* New Haven, CT: Yale University Press.

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology, 32*(5), 867-873.

Chun, D. M. (1994). Using computer networking to facilitate the acquisition of interactive competence. *System, 22*(1), 17-31.

Clark, H. H. (1979). Responding to indirect speech acts. *Cognitive Psychology, 4,* 430-477.

Conoelos, T., & Olivia, M. (1993). Using computer networks to enhance foreign language/culture education. *Foreign Language Annals, 26*(4), 527-533.

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research, 20*(6), 445-464.

Daneman M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour, 19,* 450-466.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25,* 1-18.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18,* 91-126.

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the "A" in SLA. *Studies in Second Language Acquisition, 19,* 145–171.

Ellis, N. C., & Sinclair, S. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology, 49A,* 234–250.

Engle, R., Kane, M., & Tuholski, S. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge, England: Cambridge University Press.

Fortkamp, M. B. M. (1999). Working memory capacity and aspects of L2 speech production. *Communication & Cognition, 32*(3/4), 259-296.

Freiermuth, M. R. (1998). Using a chat program to promote group equity. *CAELL Journal, 8*(2), 16-24.

Freiermuth, M. R. (2001). Native speakers or non-native speakers: Who has the floor? Online and face-to-face interaction in culturally mixed small groups. *Computer Assisted Language Learning, 14,* 169-199

Geva, E., & Ryan, E. B. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning, 43,* 5-42.

Hagen, L. K. (1990). Logic, linguistics, and proficiency testing. *ADFL Bulletin, 21*(2), 46-51.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*(1), 25-38.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Kelm, O. R. (1992). The use of synchronous computer networks in second language instruction: A preliminary study. *Foreign Language Annals, 25*(5), 441-454.

Kenyon, D. M., & Malabonga, V. (2001). Comparing examinee attitudes toward computer-assisted and other oral proficiency assessments. *Language Learning & Technology, 5*(2), 60-83. Retrieved May 20, 2005, from http://llt.msu.edu/vol5num2/kenyon/

Kern, R. (1995). Restructuring classroom interaction with networked computers: Effects on quality and characteristics of language production. *Modern Language Journal, 79*(4), 457-476.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory Language, 30*(5), 580-602.

Kost, C. (2004). An investigation of the effects of synchronous computer-mediated communication (CMC) on interlanguage development in beginning learners of German: Accuracy, proficiency, and communication strategies. Unpublished doctoral dissertation, University of Arizona.

Lantolf, J. P., & Frawley, W. (1985). Oral proficiency testing: A critical analysis. *Modern Language Journal, 69*(4), 337-345.

Lantolf, J. P., & Frawley, W. (1988). Proficiency: Understanding the construct. *Studies in Second Language Acquisition, 10*(2), 181-195.

Lantolf, J. P., & Frawley, W. (1992). Rejecting the OPI--again: A response to Hagen. *ADFL Bulletin, 23*(2), 34-37.

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

Levelt, W.J.M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology, 14,* 78-106.

Liskin-Gasparro, J. E. (1984). The ACTFL proficiency guidelines: Gateway to testing and curriculum. *Foreign Language Annals, 17*(5), 475-489.

Locke, J. (1993). *The child's path to spoken language.* Cambridge, MA: Harvard University Press.

Locke, J. (1995). Development of the capacity for spoken language. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 278-302). Oxford, England: Blackwell.

Locke, J. (1997). A theory of neurolinguistic development. *Brain and Language, 58,* 265-326.

McCarthy, M. (1998). *Spoken language and applied linguistics.* Cambridge, England: Cambridge University Press.

MacDonald, M. C., Just, M. A., & Carpenter, P. A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology, 24,* 56-98.

Medina, C. B. (1997). *Nuevos destinos.* Boston: McGraw-Hill.

Nicholas, M. A., & Toporski, N. (1993). Developing "The Critic's Corner": Computer assisted language learning for upper-level Russian students. *Foreign Language Annals, 26*(4), 469-478.

Norris, J. M. (2001). Concerns with computerized adaptive oral proficiency assessment. *Language Learning & Technology, 5*(2), pp. 99-105

Norris, J. M., & Ortega, L. (2003). Defining and measuring SLA. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 717-761). Malden, MA: Blackwell.

Omaggio, A. C. (1983). Methodology in transition: The new focus on proficiency. *Modern Language Journal, 67*(4), 330-341.

Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through sychronous CMC: Output, working memory, and interlanguage development. *CALICO Journal, 20*(1), 7-32. Retrieved March 1, 2005, from http://php.scripts.psu.edu/faculty/j/s/jsp17/articles/calico2002/

Papagno, C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30,* 331-347.

Papagno, C., Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology, 48A,* 98-107.

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Eds.), *Cognition and second language instruction* (pp. 287-318). Cambridge, England: Cambridge University Press.

Salaberry, R. (2000). L2 morphosyntactic development in text-based computer communication. *Computer Assisted Language Learning, 13*(1), 5-27.

Schenkein, J. (1980). A taxonomy for repeating action sequences in natural conversation. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 21-47). London: Academic Press.

Sempé, J., & Goscinny, R. (1997). *Les récrés du Petit Nicolas* [Young Nicholas' school breaks.] Paris: Gallimard Jeunesse.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology, 45A*(1), 21-50.

Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics, 16,* 155-172.

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

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication, 20*(1), 1-28.

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge, England: Cambridge University Press.