AN ONLINE TASK-BASED LANGUAGE LEARNING ENVIRONMENT: IS IT BETTER FOR ADVANCED- OR INTERMEDIATE-LEVEL SECOND LANGUAGE LEARNERS?

Abdurrahman Arslanyilmaz
Computer Science & Information Systems
College of STEM
Youngstown State University
United States
aarsalnyilmaz@ysu.edu

ABSTRACT

This study investigates the relationship of language proficiency to language production and negotiation of meaning that non-native speakers (NNSs) produced in an online task-based language learning (TBLL) environment. Fourteen NNS-NNS dyads collaboratively completed four communicative tasks, using an online TBLL environment specifically designed for this study and a chat tool in WebCT-Vista. Seven dyads were at intermediate-level language proficiency and the remaining seven dyads were at advanced-level language proficiency. Language production was investigated in terms of fluency, accuracy, and complexity including lexical and syntactic complexity, and negotiation of meaning was examined using the ratio of negotiated turns to total turns. The data from the chat-scripts showed that intermediate-level NNSs get involved in more negotiation of meaning than advanced-level NNSs, and advanced-level NNSs produced more accurate language than intermediate-level NNSs.

INTRODUCTION

Negotiation of Meaning and Language Production for SLA

Understanding messages and receiving comprehensible input is the only way and essential ingredient for second language acquisition (Krashen, 1985), and a considerable evidence for a causal relationship between comprehensible input and second language acquisition was reported by Long (1983). However, comprehensible input should not be conceived as simplified input adjusted to non native speakers; instead, the input should be modified by interactants in conversation through negotiation of meaning so as to ensure that it is modified to exactly the level of comprehensibility students can manage (Long, 1983; Long and Porter, 1985). In addition to negotiation of meaning, language production has also been reported to serve important functions in second language acquisition that complement the role of input (Shintani, 2011; Swain, 1985). Language production especially within a meaningful context and through interaction has been demonstrated to assist second language acquisition (Ellis and He, 1999; Morgan-Short and Bowden, 2006).

TBLL for Negotiation of Meaning and Language Production

Task-based language learning (TBLL) method promotes negotiation of meaning using 'meaning-focused' and 'communicative' task completion activities (Doughty & Long, 2003, Willis, 1996) and gives students the chance to practice language extensively within a meaningful context, engaging students in the 'authentic,' pragmatic, and contextual production of language (Doughty & Long), where language production is not the aim but 'the vehicle for attending task goals' (Willis, p.25). Negotiation of meaning and language production through negotiated interaction and meaningful task-completion activities facilitate language acquisition (Anderson, 2000) and promote second language learning (Izadpanah, 2010).

Online TBLL

The advent of information technology has improved the quality of many scientific disciplines, including language education (Chapelle, 2001; Gonzolez-Lloret, 2003), and language learning through technology has become a fact of life (Chapelle, 2001). Computer-based activities for language learning offer capabilities that cannot be easily substituted by any other language teaching procedure within a traditional classroom environment. Some of these include rapid global access at any time from any computer with Internet access, integration of graphics, audio, and text; and ease and low cost of publication (Kern and Warschauer, 2000). These capabilities make it possible to create an optimal task-based language learning environment, which is interactive, motivating, highly contextualized (or real-world like), task-oriented, and authentic. Therefore, online task-based language learning with rich multimedia experience is increasingly being used (Doughty and Long, 2003), and has been reported to be effective by many studies (Arslanyilmaz, 2010; Blake, 2000; Freirmuth and Jarrell, 2006; Gonzolez-Lloret, 2003; Keller-Lally, 2006; Smith, 2003).

Previous Studies

Because of the aforementioned benefits of TBLL to second language learning, growing body of empirical research has been conducted in this field of interest. The primary goal of these researches has been to describe, analyze, and predict the factors influencing the amount of negotiation of meaning and language production that students engage in through TBLL. Some of the variables that have been reported by these researches focused on the task-type (Blake, 2000; Smith, 2003), feedback during task-completion (Iwasaki and Oliver, 2003; Pellettieri, 2000), learners' state anxiety (Baralt and Gurzynski-Weiss, 2011), psycholinguistic considerations (Doughty & Long, 2003), format of communication - synchronous vs. asynchronous - (Kötter, 2003; Sotillo, 2000), interaction with native speakers (Tudini, 2003), similar-task videos (Arslanyilmaz, 2010), and delivery method of instructions –online vs. face-to-face- (Chen, Belkada, and Okamoto, 2004; Chun, 1994; Gonzalez-Lloret, 2003). In addition to these variables, "it is important to investigate how proficiency may impact the quality and quantity of learner interaction so as to provide the optimal opportunity for learning." (Iwashita, 2001, p. 270) Thereupon, a few studies have explored language proficiency as another variable that may have influence negotiation of meaning and language production in TBLL (Belz, 2002; Collentine, 2009; Iwashita, 2001; Pica et al., 1996; Porter, 1986); however, the number of researches conducted in this area has been insufficient to make conclusive remarks and the results of these studies have been contradictory.

The few studies with respect to the proficiency level at which students get involved in more negotiation of meaning have been contradictory. Belz (2002) conducted an asynchronous network-based foreign language study between German and American students learning English and German respectively. Belz's study reported that when high level proficient students paired with high level proficient students, they felt more comfortable and got involved in deeper discussions and participation. Pica et al. (1989, 1996, cited in Iwahsita, 2001, p. 270) stated that when a low level learner is paired up with another low level subject, compared with dyads of higher level dyads, low level learner may not know what to ask for, and repeat part of a prior utterance which they did not understand. In addition Collentine (2009) found that learners at advanced-level proficiency as compared to the ones at intermediate-level proficiency were getting involved in more interactions and producing more language because the tasks were more difficult for intermediate-level dyads and they had more knowledge gaps. On the other hand, Iwashita (2001) found that high-high dyads produced less confirmation checks and clarification request than low-low dyads even though the difference was not significant. Iwashita stated that "tasks may not be challenging enough for learners who are more proficient; and hence, they may complete the tasks without much need for negotiation of meaning."

Studies about the effect of language proficiency on language production have been insufficient and inconclusive. In one study, Porter (1986) compared the language produced by students at two proficiency level during communicative and task-based activities. She reported the amount (number of words), accuracy, and fluency of speech among other variables. Six of the learners participated in her study were advanced and six of them were intermediate. Participants completed three communicative tasks. She found that the total number of words, the fluency of language, and accuracy of language produced by advanced learners and intermediate learners were not statistically significant. Therefore, she reported that there was no clear advantage of one level over another. In another study, Iwashita (2001) reported that there was no significant difference between the amount of c-units, a measure of language quantity, produced by high-high proficient dyads and low-low proficient dyads, which was contradictory to the results reported by Porter, who stated that advance-level dyads produced more total words than intermediate-level dyads.

Research Questions
The purpose of this study is to explore whether non-native speakers at advanced- or intermediate-level proficiency produce better language and get involved in more negotiation of meaning. The research questions for this study is as follows,

1. Do advanced-advanced dyads produce more fluent language than intermediate-intermediate dyads?
2. Do advanced-advanced dyads produce more accurate language than intermediate-intermediate dyads?
3. Do advanced-advanced dyads produce lexically more complex language than intermediate-intermediate dyads?
4. Do advanced-advanced dyads produce syntactically more complex language than intermediate-intermediate dyads?
5. Do advanced-advanced dyads get involved in more negotiation of meaning than intermediate-intermediate dyads?

METHODS
Participants

Participants in this study were 14 non-native intermediate- and 14 non-native advanced-level students in an English language institute in the southern United States during the fall semester of 2006. The intermediate-level students were recruited from two sections of an intermediate-level composition course, and had been taking other intermediate-level English courses for the previous three months. The advanced-level students were randomly recruited from all advanced-level English courses offered in the institute. The advanced- and intermediate-level students represented a variety of first language backgrounds, including Korean, Mandarin, Arabic, Spanish, and Japanese. They ranged in age from 18 to 29, with the majority in their early twenties. The participants were placed in the intermediate- and advanced-level composition courses by the institute at the beginning of the semester based on a combination of their scores on TOEFL (Test of English Foreign Language), ELPE (English Language Proficiency Exam administered by the University), two in-house assessments consisting of an interview with the director of the institute and a composition test (K. Clark, personal communication, November 7, 2006).

Online TBLL Environment
An online TBLL environment was developed for this research study (see Figure 1). The environment was designed to present four tasks for students to complete in dyads. "Your Task" button was used to display the instructions for each task, and "Similar Tasks" button was used to display subtitled videos. The dyads used the chat tool provided by WebCT-Vista to complete the assigned tasks.

Figure 1: Main framework of the online TBLL environment and subtitled similar task videos for the compare the maps task.

Tasks. The online TBLL environment was designed based on task-based language learning approach focusing on the successful use of language within context rather than learning the language as an end in itself through meaningful and communicative task-completion activities. The online TBLL was designed for four tasks: two split-information and two shared-information tasks, which were developed based on the task typology by Pica et al. upon founding it to be the most informative and clear among the other task typologies. In the split-information tasks, "Compare the Maps" and "Christmas Break Trip," each student in a dyad has one part of the complete information, requires communicating to the other student in the same dyad in order to exchange the information, both student have convergent goals toward task completion, and the tasks have more than one possible outcome. In the shared information tasks, "Gifts for a Family" and "Garage Sale," two students in a dyad have the same information body available, do not have to exchange information, have convergent goals, the tasks have more than one possible outcome, and task completion involves decision making, personal preference, feeling, and attitude. These tasks were designed and developed to provide the means for students to practice language in real-life-like context while engaging them in authentic, pragmatic, contextual, and functional use of language. Students were expected to learn the language while completing the tasks, during which engaging in increased negotiation of meaning, which would free up attentional and memory resources of

students from language form and meaning resulting in paying more attention to negotiation of meaning and formulating language needed to express ideas during task completion.

In "Compare the Maps" split-information task (see Figure 2), both students in a dyad are provided with the same map containing 15 buildings along with trees, roads, and vehicles. Six of the buildings are clickable. Upon clicking one of them, one activity is displayed within the building. Three of the displayed activities are the same for both members of the dyads, and three of them are different. Descriptions of the same activities are (1) a person repairing his TV, (2) a lady studying, and (3) a child feeding her dog. Descriptions of different activities are (1) a child playing with a toy versus another toy, (2) a lady shopping for clothes versus another lady shopping for notebooks, (3) two teams playing basketball versus three people running. Dyads were asked to identify the similarities and differences between the activities.
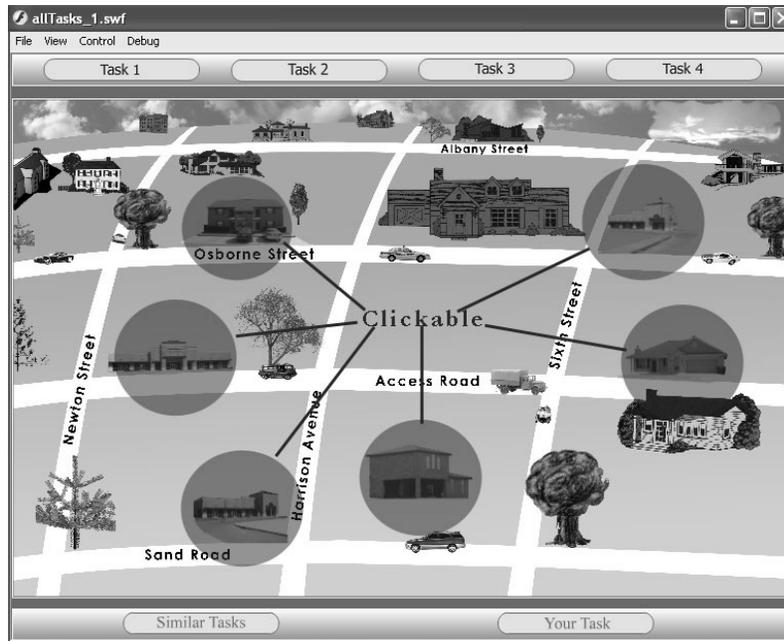

Figure 2: Compare the maps task.

In "Christmas Break Trip" split-information task (see Figure 3), dyads were asked to imagine that they have decided to take a trip during Christmas break. Each member of the dyads was provided with information about attractions, hotels, activities, and flights to three different cities. Dyads were asked to exchange information and decide which city to visit during Christmas break.

Figure 3: Christmas break trip task

In the "Gifts for a Family" shared-information task (see Figure 4), students in dyads were asked to discuss on what gift(s) to buy for each member of a family of four with whom they would be staying in the U.S. When students clicked on the house image marked with an arrow, a picture for each family member and their hobbies were displayed in the main content area. Students were then asked to discuss and decide on the amount of money to spend for each gift and what to buy for each family member based on his/her hobbies.
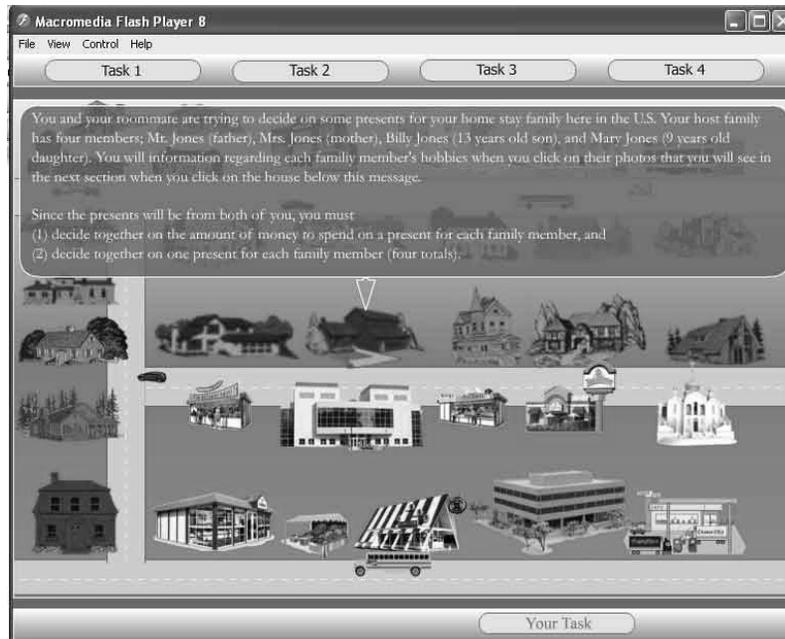

Figure 4: Gifts for a Family Task.

In the "Garage Sale" shared-information task (see Figure 5), students in the dyads were asked to imagine that they were roommates in a dormitory. Dyads are presented with their rooms and items in the room. When students click on the items, they are zoomed in and detailed information is given in the main content area. Students are asked to discuss and donate four items in their dorm room to sell at a garage sale in order to help their class raise money for a trip to Niagara Falls after talking about usefulness, value, condition, and transportation of the items, and how they would convince people to buy them.

Figure 5: Garage Sale Task.

Subtitled Videos. Each subtitled video presents a short dialog between two native speakers engaged in a task similar to – but not the same as – the one students are about to complete, see [Table 1] for brief descriptions. The videos were recorded in real-life settings, and subtitles were provided in the videos. The primary function of the videos were not to demonstrate the workings of language, nor to teach linguistic structure, but rather, as the authentic language use defined by Breen (1985), 'communicative,' 'to share experiences' (p.62) of native speakers in similar task situations with the ones the students were expected to complete. In addition, these videos were provided to the students so that they could create mental models by observing native speakers doing similar tasks in order to compare their own experiences to them to be later used in task phase. By comparing their mental models to native speakers', students would be able to fill gaps in their own linguistic resources. Furthermore, observing the videos would help students reduce cognitive processing load, facilitate conversational development, and make things less threatening (Skehan, 1998). Students who watch native speakers completing similar tasks to their assigned tasks would be motivated to pay attention to form so as to produce language like native speakers later in the task completion phase.

Table 1: *Description of Subtitled Videos*

Assigned task	Quantity	Topics of the subtitled videos
Compare the Maps	6	These focus on demonstrating and modeling an activity that is currently happening. These activities include playing a guitar, studying for an exam, making an omelet, getting ready to go home on a bike, and asking directions.
Gifts for a Family	2	Two native speakers play a couple who are deciding on gifts for their relatives, whom they are planning to visit.
Christmas Break Trip	3	Two native speakers play a couple who are making travel plans for a Thanksgiving trip.
Garage Sale	6	Speakers discuss items in their house in order to decide which to sell at a yard sale to raise money to save an endangered animal species. They talk about the value of each item, its condition, the use of each item, and reasons customers might buy them.

Treatment Conditions
Students at intermediate-level proficiency were assigned to the intermediate-level group (ILG), and students at advanced-level proficiency were assigned to the advanced-level group (ALG). As a safeguard against outliers or unbalanced groups, the director of the institute reviewed group membership and determined students belonged to their respective groups. Because students were not randomly assigned to ILG and ALG groups, but all students in the intermediate-level proficiency was assigned to ILG and all students in the advanced-level was assigned to ALG, this study should be characterized as a quasi-experimental rather than a strict experiment

design. After the formation of the groups, students in each group were randomly paired to form dyads. There were 14 dyads in the study, seven in each group. Both groups were provided with the online TBLL environment, and dyads in both groups were asked to complete the same four tasks. Both groups had access to the subtitled videos.

Procedures
All dyads met twice, each of which lasted about two hours. While the dyads in the ILG met during regularly scheduled class meetings, the dyads in the ALG met over a weekend. Two computer labs were used for the study, and one member of each dyad was assigned to work in each computer lab to ensure that none of the dyads worked face-to-face. Dyads completed two tasks in each session. Before beginning the experiment, all students were given 10 minutes of instruction for each task. After the training, each student was sent to the randomly assigned computer lab. Students were given 50 minutes to complete each task.

After learning what they are supposed to do and what components the online TBLL environment consists of, students felt the need to be prepared for the task-completion phase. Although the students were not instructed to go through a specific sequence before starting to complete their tasks, many of them started with watching the subtitled videos, where two native speakers completed similar tasks to the ones that the students were assigned to complete. While watching these videos, students paid attention to the language used by the native speakers when completing the similar tasks, which helped students to activate the task related words, phrases, and form of the second language and bring them into the working memory so as not to require much cognitive resources to access them later in the task completion phase, and helped them to acquire necessary knowledge in order to use later in the task-completion phase. In addition, watching these videos also helped students understand the goals of the tasks, what they are required to do, and nature of the outcome they would reach. Other than the subtitled videos, students explored the information in the pre-task section related to the tasks that they are required to complete. For the "Compare the Maps" task, students clicked on each one of the buildings on the interactive maps provided to them in order to see what was going on in each one of the buildings and describe these events to their partners to be able to identify similarities and differences between the activities happening in these buildings. They also paid attention to the location of the buildings including the streets where they are located at so as to be able to tell their partners which specific buildings they were referring to. For the "Christmas Break Trip" task, students were provided with information about cities that they may want to visit during the break. Students read about these cities including the fee they would pay for the hotel, airfare, food, and saw photos of the attractions they would be able to visit and see in these cities. For the "Garage Sale" task, students clicked on each one of the items in their imaginary dormitory room paying attention to the functionalities, benefits, values, worth, conditions, dimensions of these items. They also noted where the item is located at within the room, which helped them when referring to the items within the room to their partners. For the "Gifts for a Family" tasks, students clicked on each one of the family member to learn about the hobbies of each family member in order to decide what gift would be perfect for each family member. Going through the information provided to the students in the pre-task phase eased the process load that students faced during task phase, and students were able to devote more cognitive resources for task completion activities.

Data Source
The chat tool that students used to complete the assigned tasks created a transcript of their written interaction. This transcript captured all language they produced during completion of the tasks, which was analyzed for comparison between the two groups.

Data Coding
Fluency. The fluency was measured by the number of words per minute (WPM) (Kellogg, 1996) produced in written language using the chat tool. The total number of words produced by each dyad in both groups was tallied and divided by 200, the total number of minutes spent to complete the four tasks. When counting the words, any orthographic unit bound by spaces, and including proper nouns and acronyms (e.g. TAMU) was counted as one word. 'Concatenated forms (e.g. gonna, wanna, sorta) and contracted forms (I'm, we're)' (Andersen & Johnson, 1973, p. 151) were counted as one word. Hyphenated words (quarter-mile) were counted as two words. Back-channeling cues (e.g. hhhhh hh, um, oh, hey) were not counted as words. Repetitions resulting from an attempt to correct a lexical error were counted as one word (e.g., A:hotel pee and else, A:fee, not pee). To control the influence of computer proficiency on the results of fluency, before starting the experiment, all students were asked to fill out a background questionnaire designed to gather information about students' computer proficiency. The results of an independent samples t test with the computer proficiency scores as the dependent variable and the groups as the independent variable shows that students in ILG were not significantly different than students in ALG in terms of their computer proficiencies (p = .61).

Syntactic Complexity. Syntactic complexity was measured by the number of subordinate clauses per 'C-unit' (Robinson, 2001). Following the definition by Rulon and McCreary (1986), any 'word, phrase, or sentence' that in some way contributed 'pragmatic or semantic meaning' to a conversation regardless of grammaticality were counted as a C-unit (e.g., Mine is a old woman (one C-unit), not a girl (one C-unit)). Elliptical answers to questions were counted as one C-unit (e.g., A:"Where are you going? (one C-unit)," B:"to the store (one C-unit)). Some words were included with the following or previous C-units and counted as one C-unit if they seemed to form one semantic unit (e.g., "Now (not a separate C-unit). I stop temporary. (one C-unit)"). Repetitions in an attempt to correct the previous C-units by the same speaker were not counted as separate C-units (e.g., "I know that on the fist school (one C-unit). First school (not a separate C-unit)"). When the speaker's turn is interrupted by the listener, both halves of the speaker's turn was counted as one C-unit if they are one clause (e.g. A:"like Walt Disney World (one C-unit)" B:"I like Walt Disney too" A:"and Discovery Cove (not a separate C-unit)"). When a speaker used more than one line for one communication unit, they were counted as one. If a single sentence was formed as a result of a combination of two simple sentences with a conjunction, they were coded as two separate C-units (e.g., "Yes, map is the same (one C-unit) but people inside are different (one C-unit)"). Sentences with correlative conjunctions were coded as one C-unit if they had a compound subject, object, adjective, or adverb (e.g., "it is close to many things and to the airport" (one C-unit)). However, if they linked two verbs, then they were coded as two different C-units (e.g., "a woman shopping (one C-unit) and buying a T-shirt" (one C-unit)).

The subordinate clauses reduced to a phrase by omitting the relative pronoun and the part of the verb were counted as one subordinate clause (e.g., "We are completely the same in the building with two cars"). If the subordinate was used as an answer to a question by the other student, it was not counted (e.g., A:"Why?" B:"because the children like music (no subordinate clause but one C-unit).

Lexical Complexity. Lexical complexity was measured by mean segmental type token ratio (MSTTR), 'an index that appears to have been originally recommended by Johnson' (Malvern & Richards, 2002, p. 88). To find out MSTTR, the students' written language products were divided into segments of 200 words, the type-token ratio (TTR) of each segment was calculated, and the average of the type-token ratios for the segments were taken. MSTTR was used because 'TTR is a function of sample size: larger samples of words will give a lower TTR' (Malvern & Richards, 2002), which has also been criticized by many other scholars. MSTTRs were calculated by the WordSmith tool (Scott, 2008). When calculating MSTTRs, this tool lists all unique words with the number of times they were used in the text and the context in which they were used. Each of these unique words was analyzed in context to see whether it was a real word because the tool considers every utterance bound by space as a real word. For example, some utterances (e.g., "T," "LL"), back channeling cues (e.g., "HHH," "HAHHAHA"), and typos within separate occurrences of the same word without typing error (e.g., "YOU^^") are considered as a separate word by the tool. These utterances were taken out from the list of unique words.

Accuracy. Accuracy was measured by the ratio of 'error-free T-units' to total 'T-units' (Bygate, 2001). A 'T-unit (terminable unit)' defined as an independent clause and subordinate clauses 'attached to or embedded in it' (Kern, 1995). The difference between T-units and C-units is that T-units are easily and objectively identifiable by identifying independent clauses and subordinate clauses attached to them. On the other hand, identifying C-units are more subjective because a C-unit may not be an independent clause or subordinate clause as long as it somehow contributes pragmatic or semantic meaning to a conversation (e.g. "to the store" as an answer to a question "Where are you going?" is a C-unit but not a T-unit), and this contribution can be subjective depending on how it is perceived by the rater based on the context in which the word or phrase is used (e.g., "I like Walt Disney World but not Discovery Cove" could be coded as one C-unit by one rater but as two C-units by another rater depending on whether they perceive the sentence contributes one meaning or two different meanings to the conversation).

The T-units and errors in the T-units were calculated based on the combination of two guidelines, one of which was used by Sotillo (2000) to calculate accuracy in synchronous and asynchronous communication and the other one was used by Polio (1997) to compare three different measures for linguistics accuracy (see the Appendix for error guidelines). The two guidelines are very similar to each other with trivial differences.

Ratio of negotiated turns to total turns. The negotiation of meaning was measured by the ratio of negotiated turns to total number of turns, which is chosen to remove the effect of the amount of talk on the amount of negotiation of meaning. In order to calculate the ratio of negotiated turns to total turns, it was first necessary to code the data and count the number of negotiation of meaning sequences (negotiated turns).

All negotiation of meaning sequences were identified using the model developed by Varonis and Gass (1985) and revised by Smith (2003). As defined in the model, negotiation of meaning sequences consists of two parts: trigger and resolution. The trigger <T> is the utterance or portion of an utterance on the part of the speaker that results in some indication of non-understanding on the part of the listener (Varonis & Gass, 1985, p. 74). Five types of triggers were found in this study: lexical (see Excerpt 1), content (see Excerpt 2), syntactic (see Excerpt 3), task complexity (see Excerpt 4), and discourse (see Excerpt 5).

The resolution part of a negotiation of meaning sequence consists of an indicator, and perhaps a response, the reaction to the response, a confirmation, and a reconfirmation (Smith, 2003; Varonis & Gass, 1985). Indicators, the written communication where the listener signals that there is a non-understanding, were coded <I>. Responses, where the original speaker attempted to clear up the non-understanding, were coded <R>. Reactions to the response in which the listener signaled a degree of understanding were coded <RR+> and those that indicated continued difficulty with the speaker's response were coded <RR->. Confirmations, which indicate a positive reaction to the response <RR+>, that is, that some degree of understanding was achieved by the listener, were coded <C>. Reconfirmations, where even a minimal response to the respondent's confirmation occurred were coded <RC>.

Excerpt 1
H.Y.: shall we but it?   <T>
A.Q.A: what shall
A.Q.A.: what is shall   <I>
H.Y.: shall we go to buy it?   <R>

Excerpt 2
A.H.: I was looking for the attractions and that just make
me be excited   <T>
L.Y.: What kind of attractions? Could you tell me?   <I>
A.H.: There we can visit Walt Disney World, Discovery
Cove,Epcot Center and the Universal Studios
A.H.: it could be fun   <R>
L.Y.: Yes, that sounds very good. How about the prices?   <RR+>Implicit<TAR>

Excerpt 3
A.Q.A.: what do you thing about Brushed that machian do
damage the papers   <T>
H.Y.: whats mean?   <I>
A.Q.A.: machian do damage the papers   <R>
H.Y.: what damage?   <RR->
A.Q.A.: it is next the referajrater   <R2>
H.Y.: ok   <RR+>

Excerpt 4
S.K.: In book store building, a woman is looking around the shop,
She has an aggi T- shirt.   <T>
C.D.: In my picture she has a black coat
C.D.: Has she a back bag?   <I>
S.K.: No, she wears yellow coat.
S.K.: She does not have any bag.   <R>
C.D.: Ok. They are different.   <RR+>Explicit
S.K.: OK.   <C>

Excerpt 5
I.J.: do you know tori?   <T>
K.K.: tori   <I>
I.J.: she is in the video   <R>
K.K.: tori is a couch of Newyork Yankees
K.K.: coach   <RR->Implicit<TD>

Negotiation of meaning sequences consisting of a T-I-R, a T-I-R-RR, a T-I-R-RR-C, or a T-I-R-RR-C-RC were identified and counted, and the ones consisting of only a T-I were not included in the data analysis. After

identifying the negotiation of meaning sequences, the total number of turns and negotiated turns were counted. A turn is considered to occur whenever there is a transfer of the floor from one student to another. Negotiated turns are those turns that occur in negotiation of meaning sequences, the most widely used model of negotiation (Smith 2003, p. 39). Finally, a ratio of negotiated turns to total turns was calculated to find out the amount of negotiation of meaning occurred during completion of the tasks.

Reliability

A randomly selected 10% of the written interaction produced by dyads in ILG and ALG during completion of tasks was coded by an independent rater using the same procedures described in this study for identifying subordinate clauses, C-units, T-units, error-free T-units, and negotiation of meaning sequences. These subordinate clauses, C-units, T-units, error-free T-units, and negotiation of meaning sequences identified by the independent rater were compared to the ones identified by the researcher. The researcher and the independent raters examined the coding differences, and discussed these differences for clarification and making sure that they both coded the units with the same understanding of the criteria for the coding. They agreed on the coding for the 89% of the subordinate clauses, 91% of the C-units, 93% of the T-units, and 90% of the error-free T-units for the language produced by dyads in the ALG. They also agreed on the coding for the 90% of subordinate clauses, 93% of the C-units, 92% of the T-units, and 90% of the error-free T-units for the language produced by dyads in ILG.

Another random selection of 10% of the written language produced by dyads in ILG and ALG was coded by an independent rater to identify negotiation of meaning sequences. The independent rater used the same procedures as described in this study in identifying the negotiation of meaning sequences. After coding the negotiation of meaning sequences including total number of turns and the number of negotiated turns occurring within the negotiation of meaning sequences, the independent rater and the researcher came together to discuss how each one of them coded. They compared the coding by both of them, and clarified any misunderstandings in the way each one them coded. As a result, they agreed 88% of the negotiated turns for the language produced by the ALG and 89% of the negotiated turns for the language produced by ILG. The two rater agreed 100% of the total turns.

Data Analysis Techniques

Because language production and input comprehension are two different measures, two independent statistical analyses were executed to determine the effect of language proficiency on language production and negotiation of meaning. Analysis of multivariance (MANOVA) was used to analyze the effect of language proficiency on language production because the experiment affects all four dependent variables (fluency, syntactic complexity, lexical complexity, and accuracy) separately and in combination with each other. The level of significance (alpha) was set at.05. Hotelling T^2 was used to determine overall multivariate significance of dependent variables on the groups.

In order to analyze the effect of language proficiency on negotiation of meaning, this study used only one 2-tailed independent group t-test. The ratio of negotiated turns to total turns produced by ALG was compared to the ones produced by ILG, and these two sets of dependent variables are not related to each other in any way because the negotiated and total turns in the two groups are produced by different students in two different groups.

RESULTS AND DISCUSSIONS

Language Production

By looking at [Table 2], we are able to determine the relative amount of language that was produced while students were engaged in task-completion activities. In terms of fluency, students in ALG produced almost 1.42 times more words than students in ILG during the same time span. In terms of lexical complexity, students in the ALG produced almost 1.41 times more unique words, and 1.04 times more lexically complex language as measured by mean segmental type token ratio (MSTTR) than students in the ILG. In terms of syntactic complexity, surprisingly, students in the ILG produced about 1.22 times more subordinate clauses per C-unit than students in the ALG. In terms of accuracy, the ratio of error free T-units to total T-units were 1.59 times better with the students in the ALG than with the students in the ILG.

Table 2: *Total Number of Words, Unique Words, C-Units, Subordinates, and Total and Error-free T-units Produced across Groups*

Groups	Total words	Unique words	C-units	Subordinates	T-Units	SPC	MSTTR	WPM	EFT

					Total	Error-free				
ILG	11800	4312	2193	329	1477	644	0.61	51.96	33.71	0.41
ALG	16764	6119	2920	359	2138	1401	0.50	54.44	47.9	0.65

Note: SPC= Subordinate clause per communication unit; MSTTR= Mean segmental type token ratio; WPM= Word per minute; EFT= Ratio of error-free T-units to total T-units

A multivariate analysis of variance was conducted to assess if there were differences between the two groups (ILG, ALG) on four language production measures (words per minute, subordinate clause per C-unit, mean segmental type token ratio, and ratio of error-free T-units to total T-units) considered together. MANOVA multivariate test output as seen in [Table 3] shows a significant multivariate main effect for Group at .05 significance level, Hotelling's T = 5.22, F = 11.75, p=.001. This result suggests that students in the ALG based on the composite of the four dependent variables significantly differ from students in the ILG. Partial eta squared value of .84 reported in [Table 3] indicates that 84 percent of the variance in language output (the new combined dependent variable) can be accounted for by the two groups at two different proficiency level (the independent variable). Therefore, a large proportion of the variance, .84, is explained by the between groups SSCP matrix, or in other words by the experiment.

Table 3: *Comparison of Mean Number of WPM, SPC, MSTTR, and EFT across Groups*

Groups	Value	F	P	Observed power	Partial eta squared
Hotelling's T	5.22	11.75	.001[*]	.993	.84

Note: WPM = Word per minute, SPC = Subordinate clause per minute, MSTTR = Mean segmental type token ratio, EFT = Ratio of error-free T-units to total T-units[*] p < .05

Examination of the coefficients for the linear combinations distinguishing the two groups indicated that the ratio of error-free T-units to total T-units (F(1,12) = 9.93, Partial eta-squared = .45, p = .008), word per minute (F(1,12) = 5.33, Partial eta-squared = .3, p = .04), and mean segmental type token ratio (F(1,12) = 6.2, Partial eta-squared = .34, p = .028) contributed significantly. However, subordinate clause per communication unit (F(1,12) = 1.56, Partial eta-squared = .115, p = .23) did not contribute significantly to distinguish the groups. By partial eta-squared, the ratio of error-free T-units to total T-units effect appears stronger than the mean segmental type token ratio effect, which appears stronger than the word per minute effect, which appears stronger than subordinates per C-units effect.

Bonferroni correction was applied to the Univariate ANOVA test statistics by dividing 0.05 by 4 (number of dependent variables) in order to ensure against a Type I error. Hence the values need to be smaller than 0.0125 for the results to be significant. The Univariate ANOVAs indicated that ratio of error-free T-units to total T-units (p = .008) were significantly higher for students in ALG than students in ILG. Word per minute (p = .04), mean segmental type-token ratio (p = .028) were also significantly higher for students in ALG than students in ILG before the Bonferrroni correction is applied to the results but not after the Bonferroni correction is applied to the results. Subordinate clause per communication unit (p = .24) was not significantly higher for students in ILG than students in ALG both before and after the Bonferroni correction is applied to the results. The results suggest that students in advance-level proficiency group (ALG) produced significantly more correct language than student in intermediate-level proficiency group (ILG).

However, fluency and lexical complexity of the language produced by students in ALG was not statistically significant than students in ILG even though ALG students produced more fluent and lexically complex language than ILG students. In addition, although syntactic complexity of language produced by students in ILG was more complex than students in ALG, it was not statistically significant. It is therefore appropriate to report that online task-based second language learning is more beneficial to students in advanced-level course than students in intermediate-level course in the production of language in general when complexity, fluency, and accuracy are combined, and in the accuracy of language produced specifically.

Negotiation of Meaning
As seen in [Table 4], negotiated turns accounted for about 8.8% of the total turns generated by dyads in the advanced-level group (ALG). In contrast, negotiated turns accounted for about 25% of the total turns generated

by dyads in the intermediate-level group (ILG). This result suggests that students at intermediate-level language proficiency engaged in more negotiated interaction than students at advanced-level language proficiency. These figures also suggests that students at intermediate-level language proficiency produced about 5.53 times more negotiation of meaning sequences, about 8.33 times more negotiated turns, and about 2.99 times more total turns than students in the advanced-level language proficiency. The ratio of negotiated turns to total turns was about 2.8 times more with the students at intermediate-level language proficiency than with students at the advanced-level language proficiency. Therefore, intermediate-level proficiency students got involved in negotiation of meaning more often than students at advanced-level proficiency during task completion activities.

Table 4: *Total Negotiation of Meaning Sequences, Total Turns, Negotiated Turns*

Groups	Negotiation of meaning sequences	Negotiated turns	Total turns	Ratio of negotiated turns to total turns
ALG	15	42	475	0.88
ILG	83	350	1422	2.46

[Table 5] shows the results of an independent samples t-test with the percentage of turns negotiated as the dependent variable and groups as the independent variable. This table shows that dyads at intermediate-level language proficiency produced a significantly higher percentage of negotiated turns than dyads at advanced-level language proficiency. One of the reasons for the significant result with a small number of participants is the low within group variance, variance between students prior to the experiment. This was undoubtedly partly a result of the sample selection procedure in that participants were chosen from a homogenous group of English language students. Another reason for the significant result is the high between group variance, mean differences between the ratios of negotiated turns to total turns for students who were in ILG and for students who were in ALG. This high between group variance is attributed mostly by the experiment.

The answer for our research question appears to be that non-native speakers at intermediate-level language proficiency produce more negotiation of meaning than students in advance-level language proficiency in an online task-based language learning environment.

Table 5: *Comparison of Mean Percentage of Negotiated Turns to Total Turns across Groups*

Group	N (Dyads)	M	SD	T	df	Sig. (2 tailed)	99% Conf. Int. Lower	Upper
ALG	7	.10	.1	2.69	12	.02*	.024	.23
ILG	7	.23	.073					

*p<.05

CONCLUSION

This study demonstrated that students in advanced–level proficiency produce better language in general than students in intermediate-level proficiency in an online task-based language learning environment. One of the reasons for this result could be the fact that advanced-level students had more linguistic resources, task-related words, phrases, language forms in their working memory both as a result of higher level of proficiency they were at and as a result of the higher level of comprehension they were able to attain from the videos and information provided in the pre-task phase. Accordingly, they were able to access more diverse task related words, phrases, and forms into their working memory, bring them faster, and put them together more accurately. In addition, they were also able to fill more gaps in their own linguistic resource to help them produce better language when fluency, accuracy, and complexity combined. This result confirms the finding by Collentine (2009), who reported that learners at advanced-level proficiency as compared to the ones at intermediate-level proficiency produced more language. However, this result is contradictory to the result reported by Porter (1986), who claimed that "advance learners and intermediate learners bring comparable skills of interaction to their discussions and there is no clear advantage of one level over another." (Porter, 1986, pp. 212-212) One explanation for the discrepancy could be the fact that students in Porter's study completed the tasks face-to-face and the data consisted of oral language produced by students on paper-based tasks; on the other hand, students in this study completed online and interactive tasks communicating through an online chat tool. In addition, unlike the current study, participants in Porter's study were not provided with the similar task videos, which was reported to have a positive impact on the amount of language produced (Arslanyilmaz, 2010).

In addition, univariate test statistics showed that advanced-level students produced more accurate language than intermediate-level students in an online task-based language learning environment. This result is not surprising either because of the fact that students' language in advanced-level proficiency is supposed to be grammatically

more accurate than students' language in intermediate-level proficiency. Furthermore, it is possible that advanced-level students were able to pick out more language produced by native speakers in the similar task videos than intermediate-level students resulting in more native-like language with less grammatical errors. This could be because advanced-level students were able to establish more relationships between the input provided by native speakers in the similar task videos and their prior linguistic resources. However, this result is not in congruence with the result reported by Porter (1986), who reported that the accuracy of language produced by advanced-level student dyads was the same as the accuracy of the language produced by intermediate-level student dyads. As mentioned above, this could be attributed to the difference in the presentation of the tasks, online and interactive vs. paper-based, and the mode of interaction between the students during task completion, face-to-face and oral vs. online and written.

On the other hand, fluency and lexical complexity was not significantly better for students in advanced–level proficiency than students in intermediate-level proficiency. This result complies with the result reported by Porter, who stated that there was no significant difference between the fluency of language produced by the advanced- and intermediate-level student dyads, and by Iwashita (2001), who reported that there was no significant difference between the amount of c-units, a measure of language quantity, produced by high-high proficiency dyads and low-low proficiency dyads. This could be due to the limited number of participants involved in the study, which was also the case in Porter's study with a total of six participants in each proficiency level. In the current study, students at advanced-level produced 1.42 times more words and 1.41 times more lexically complex language than students at intermediate-level, and these were not enough to warrant a significant result, which could have been possible if there were more students participated in the study.

One interesting result of this study was that students in intermediate-level proficiency produced syntactically more complex language than students in advanced-level proficiency although it was not statistically significant. One possible explanation of this could be that intermediate-level students produced less language within the same time frame than advanced-level students. Therefore, students in intermediate-level proficiency took more time in constructing their sentences than students in advanced-level proficiency when completing the tasks.

The results showed that students in intermediate-level proficiency group got involved in significantly more negotiation of meaning that students in the advanced-proficiency group. This result was aligned with the finding by Iwashita (2001), who also found that high-level proficient students produced less negotiated interaction than low-level proficient students. This result also supports his reasoning that less negotiation of meaning could be as a result of tasks being too easy for them, who completed them without too much negotiation of meaning. But, it is also possible to suggest that advanced-level students did not have to prevent and repair breakdowns in communication and to sustain the conversation, called interactional modifications (Porter, 1986), as much as intermediate-level students did. In another words, students at intermediate-level had more trouble with communicating with each other than students at advanced-level, which is also supported by the aforementioned results of this current study. Students at the intermediate-level had to negotiate meaning of a linguistic form, conversational structure, or message content in order to achieve a mutual understanding before they move forward with the completion of the tasks. Advanced-level students on the other hand were able to understand each other better, and therefore did not need to negotiate the meaning of a linguistic form, conversational structure, or message content. Therefore, intermediate-level students had significantly more instances of negotiated meanings until they achieve an acceptable level of understanding (Long, 1996).

The result is contradictory to the report made by Pica et al. (1989, 1996, cited in Iwahsita, 2001, p. 270), who state that low level learner paired with another low level learner may not know what to ask for; therefore, may not get involved in as much negotiation of meaning as high level learner paired with another high level learner. The result also contradicts to the study results by Collentine (2009), who reported that learners at advanced-level proficiency as compared to the ones at intermediate-level proficiency were getting involved in more interactions.

The implication of this study is that online TBLL environment should be integrated into the second language classes for both intermediate- and advanced-level second language students. While advanced-level students enhance their language production in terms of fluency, accuracy and complexity, intermediate-level students get involved in more negotiation of meaning during task completion activities. Therefore, online task-based language learning environments will help advanced-level second language learners with their language production skills, intermediate-level students will benefit from it through input comprehension.

The online TBLL environment was designed for two task types: shared- and split-information tasks, and the effect of the task type that may have contributed to the result of this study was not investigated in this study. Therefore, future research studies should be conducted to explore whether task type had any contribution in the language proficiency's effect on language production and negotiation of meaning that second language learners produced in the online TBLL environment.

There were a few limitations in this study that needs to be identified and acknowledged. One limitation was the fact that the experiment for students in ILG group and for the students in ALG group was not conducted at the same time. While the dyads in the ILG met during regularly scheduled class meetings, the dyads in the ALG met over a weekend. Although there is no evidence, it is possible that the experimentation time might have had influence over the results of the study. Another limitation was the limited number of students involved in the study. Only 28 students, 14 in each group, involved in the study, which could have been more if there were more students enrolled in the English Language Institute, where the students were recruited from.

REFERENCES

Andersen, E.S., & Johnson, C.E. (1973). Modifications in the speech of an eight-year-old as a reflection of age of listener. *Stanford Occasional Papers in Linguistics, 3*, 149-160.

Anderson, J. (2000). *Learning and memory: An integrated approach*. New York: John Wiley and Sons.

Arslanyilmaz, A. & Pedersen, S. (2010b) Improving Language Production via Subtitled Similar Task Videos. *Language Teaching Research Journal, 14*(4).

Baralt, M., & Gurzynski-Weiss, L. (2011). Comparing learners' state anxiety during task-based interaction in computer-mediated and face-to-face communication. *Language Teaching Research*, *15*, 201-29.

Belz, J. (2002). Social dimensions of telecollaborative foreign language study. *Language Learning & Technology, 6,* (1) 60-81.

Blake, R. (2000). Computer mediated communication: a window on L2 Spanish interlanguage. *Language Learning & Technology, 4*(1), 120-136. Retrieved May 26, 2003, from http://llt.msu.edu/vol4num1/blake/

Breen, M. (1985). Authenticity in the language classroom. *Applied Linguistics,* 6/1, 60-70.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogical tasks: Second language learning, teaching and testing.* Harlow: Pearson Education. 23-48.

Chen, J., Belkada, S., & Okamoto, T. (2004). How a web-based course facilitates for acquisition of language for academic purposes. *Language Learning and Technology, 8*(2), 33-49.

Chapelle, C. (2001). Computer applications in second language acquisition: Foundations for teaching, testing, and research. Cambridge: Cambridge University Press.

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

Collentine, K. (2009). Learner use of holistic language units in multimodal, task-based synchronous computer-mediated communication. Language Learning & Technology, 13, 68–87. Retrieved from http://llt.msu.edu/vol13num2/collentine.pdf

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

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*, 285-301.

Ellis, R. & Yuan, F. (2004). The effects of planning on fluency, complexity and accuracy in second language narrative writing. *Studies in Second Language Acquisition.* 26 (1). 59- 84.

Freiermuth, M., & Jarrell, D. (2006). Willingness to communicate: Can online chat help? International Journal of Applied Linguistics, 16, 189-212.

González-Lloret, M. (2003). Designing task-based CALL to promote interaction: En Busca de Esmeraldas. *Language Learning & Technology,7*(1), 86-104. Retrieved February 3, 2003, from http://llt.msu.edu/vol7num1/gonzalez/default.html

Izadapanah, S. (2010). A study on Task-Based Language Teaching: From theory to practice. US-China Foreign Language, 8(3), 47-56. Retrieved from EBSCOhost

Iwasaki, J., & Oliver, R. (2003). Chatline interaction and negative feedback. *Australian Review of Applied Linguistics, 17,* 60-73.Pellettieri, J. (2000). Negotiation in cyberspace: The role of chatting in the development of grammatical competence. In M. Warschauer & R. Kern (Eds.), *Network-based Language Teaching: Concepts and Practice* (pp. 59-86). Cambridge, England: Cambridge University Press.

Iwashita, N. (2001). "The effect of learner proficiency on corrective feedback and modified output in nonnative-nonnative interaction". System 29, 2.267-287.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. E. Ransdell (Eds.). *The Science Of Writing: Theories, Methods, Individual differences and applications* (pp. 57-71). Mahwah, NJ: Lawrence Erlbaum Associates.

Kern, R. & Warschauer, M. (2000) Theory and practice of network-based language teaching. In: Warschauer, M. & Kern, R. (eds.) Network-based language teaching: Concepts and practice. New York: Cambridge University Press, 1-19

Keller-Lally, A. (2006). Effect of task-type and group size on foreign language learner output in synchronous computer-mediated communication. Unpublished doctoral dissertation. University of Texas, Austin.

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

Kötter, M. (2003) Negotiation of Meaning and Codeswitching in Online Tandems. Language Learning and Technology, 7(2): 145–72

Krashen, S. (1985). *The input hypothesis*. London: Longman.

Long, M. (1983). Native speaker/nonnative speaker conversation in the second language classroom. In M. A. Clarke and J. Handscombe (Eds.), *On TOSEL '82* (pp. 207-225). Washington, DC: TESOL.

Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly, 19*(2), 207-28.

Malvern, D. D., & Richards, B. J. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, *19*, 85–104.

Morgan-Short, K. & Bowden, H. W. (2006). Processing Instruction and Meaningful Output-Based Instruction: Effects on Second Language Development. *Studies in Second Language Acquisition, 28*(1), 31-65.

Pellettieri, J. (2000). Negotiation in cyberspace. The role of chatting in the development of grammatical competence. In M. Warschauer, & R. Kern (Eds.), Network-based language teaching: Concepts and practice (pp. 59-86). Cambridge, England: Cambridge University Press.

Pica, T., Lincoln-Porter, F., Paninos, T., and Linnell, J. (1996). Language learners' interaction: How does it address the input, output, feedback needs of second language learners? TESOL Quarterly 30, 1:59-84.

Pica, T., Holliday, L., Lewis, N., and Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on learner. Studies in Second Language Acquisition, 11:63-90.

Pica, T. (1987). The impact of interaction on comprehension. *TESOL Quarterly 21*(4), 737-58.

Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning*, *47*, 101–143.

Porter, P. A. (1986). How learners talk to each other: Input and interaction in task-centered discussions. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 200-222). Cambridge, MA: Newbury House.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*: 27-57.

Rulon, K. A., & McCreary, J. (1986). Negotiation of content: Teacher-fronted and small group interaction. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 182-189).

Cambridge, MA: Newbury House.

Scott, M. (2008). *Wordsmith tools version 5.0*. Oxford*:* Oxford University Press, Oxford, England.

Shintani, N. (2011). "A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners." Language Teaching Research 15(2): 137.

Smith, B. (2003). Computer-mediated negotiated interaction: An expanded model. *The Modern Language Journal, 87*, 38-58.

Sotillo, S. M. (2000). Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning & Technology, 4*(1), 82-119.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input in its development. In S.M. Gass, & C.G. Madden, (Eds.), *Input in second language acquisition* (pp. 235-53). Rowley, MA: Newbury House.

Tudini, V. (2003). Using native speakers in chat. Language Learning & Technology, 7, 3, 2003, pp. 141-159. http://llt.msu.edu/vol7num3/pdf/tudini.pdf - accessed May 29, 2009.

Varonis, E. M., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics, 6* (1), 71-90.

Willis, J. (1996). *A framework for tasked-based learning.* London: Longman

Appendix: Error guidelines
1. Improper spellings of proper nouns were counted as errors (e.g. I want to go to new York city)
2. Spelling errors were not counted (e.g. The condition of thoes chairs is good)

3. Missing commas in restrictive clauses were counted (e.g. We must go to Orlando where we can visit the Walt Disney)
4. Inappropriately placed extra commas were counted (e.g. Hotel costs $72, per day)
5. Missing commas between clauses or after prepositional phrases were not counted (e.g. Because she likes to help with housework she needs to spent time with her mother in the kitchen)
6. Tense/reference errors were counted within the context of the preceding discourse (e.g. Why did not you make some omelet if you are hungry)
7. Overgeneralizations of –ed markers were counted (e.g. I did not remembered that I had a chocolate)
8. Plural –s to inappropriate contexts was counted (e.g. Have you visited Tori's houses)
9. Omissions of third person singular, -s marker, were counted (e.g. She enjoy playing with her car)
10. Article omissions were counted (e.g. I think it is not necessary us)
11. Syntax errors were counted (e.g. We do not need also the round table)
12. Morphology errors were counted (e.g. It is broken for a long time)