# LEARNER AUTONOMY IN A TASK-BASED 3D WORLD AND PRODUCTION

**Karina Collentine**

**Northern Arizona University**

This study contributes to the research on learner autonomy by examining the relationship between Little's (1991) notions of 'independent action' and 'decision-making', input, and L2 production in computer-assisted language learning (CALL). Operationalizing 'independent action' and 'decision-making' with Dam's (1995) definition that focuses on 'choice', the present study examines whether, while learners are engaged in a CALL task, their choices—termed *autonomous moves*—within a 3D environment and the subsequent input they receive predict the linguistic complexity and accuracy of their production in synchronous computer mediated communication (SCMC). A total of 58 third-year university-level learners of Spanish participated in two murder-mystery tasks coupling a 3D segment—containing embedded user-tracking features—with an SCMC segment. Four regression analyses examined the potential impact of learners' choices within the 3D environment and the input resulting from those decisions on their production. The results suggest that learners' linguistic complexity and accuracy while completing CALL-based tasks is influenced by both their autonomous moves and the linguistic characteristics of the input they receive (as a result of their autonomous moves).

**Keywords:** Computer-Assisted Language Learning, Second Language Acquisition, Task-based Instruction

## INTRODUCTION

With widespread access to technology, learners are increasingly using CALL materials in a learner-centered approach where they take control of their own learning, on their own time, and for their own purposes. These materials include virtual and 3D environments with gaming-like experiences (Darasawang & Reinders, 2010; Sykes, 2009). Highly interactive, multi-sensory environments provide access to real world simulations (Pantelidis, 1993; Schwienhorst, 2008), popularizing online multiuser virtual environments (e.g., Second Life) and massively multiplayer online games (e.g., World of Warcraft). In these autonomous learning environments entailing "independent action" and "decision-making" (Little, 1991, p. 4), it is essential that learners become cognizant of how to learn by raising their metalinguistic awareness and participating in tasks that motivate L2 communication. Fischer (2007) and Schwienhorst (2008) argue that learners in these environments should develop metacognitive abilities, strategies, and have opportunities for reflection (e.g., on input characteristics or their own learning strategies).

According to van Lier (2010), studying autonomy—and so its operationalization—with empirical rigor in controlled experimental settings has been a challenge since the construct 'learner autonomy' is multifaceted. One ramification is that autonomy research has not addressed the direct impact of autonomous learning on L2 communicative competence. Therefore, this study examines autonomous learning in a task-based 3D CALL environment, operationalizing autonomy with Dam's (1995) notion of 'choice': decisions learners make in the 3D world to interact with the environment towards solving the task. Specifically, it explores the relationship between what learners choose to interact with—termed 'autonomous moves'—in CALL, the nature of the subsequent input they receive, and the linguistic complexity and accuracy of their production.

**Learner Autonomy in CALL**

According to Little (1991), autonomy is "the learner's psychological relation to the process and content of learning—a capacity for detachment, critical reflection, decision-making, and independent action" (p. 45). Schwienhorst (2003) emphasizes that autonomy in CALL involves learners in critical self-evaluation and reflection as well as self-determination so that they take control over and responsibility for their development. Autonomous learning in CALL can help learners modify input and gauge learning, monitor their progress, and reflect on and prioritize their learning (Darasawang & Reinders, 2010; Gick, 2002; Reinders, 2006, 2007; Toogood & Pemberton, 2002).

Van Lier (2010) encourages researchers to find ways to increase autonomy research's ecological validity by bringing experimental, quantitative approaches to the conversation. Fischer (2007) argues that CALL researchers can study learner autonomy in one of two ways: (1) collecting data on students' learning strategies and degree of reflection while using CALL materials (e.g., self reports); or (2) identifying (e.g., for future CALL developers) design principles that will promote learner autonomy. The latter approach has been underutilized (cf. Fischer, 2007, p. 418) and is adopted in the present study.

In the context of this challenge, van Lier (2010) argues that autonomy is multifaceted, noting: "we have no clear, universally accepted definitions of the terms under discussion, let alone *operationalized ones* [author emphasis] that can be used to conduct rigorous empirical research" (p. xiv). Operationalization depends on the CALL materials one uses to foster learner autonomy. Much autonomy research takes a social-interactive approach (Schwienhorst, 2003). Yet, many CALL activities do not entail social interaction. For instance, in the 3D environment described below, students independently interact and gather information from objects and non-player characters (NPCs), or 3D representations of humans that users have the option of prompting for information relevant to the task (as programmed by the 3D environment's designer). Dam's (1995) definition of autonomy is particularly helpful in operationalizing the construct in non-social-interaction environments because it encourages researchers to consider learners' choices: "a learner qualifies as an autonomous learner when he independently chooses aims and purposes and sets goals; chooses materials, methods and tasks; exercises choice and purpose in organizing and carrying out the chosen tasks; and chooses criteria for evaluation" (Dam, 1995, p. 45). Admittedly, this operationalization limits the study of autonomy to focusing on Little's (1991) criteria of 'decision-making', and 'independent action'. Still, such a focus may have the benefit of contributing ecological validity to the body of research.

Since the present study examines the relationship between autonomy in CALL and learner production, the role of the input must also be considered. Input in autonomous learning affects linguistic awareness (cf. Schwienhorst, 2003), and so learner output. In this study's 3D world (as well as in many CALL applications), learners make choices about what and how much input they access. Such input, thus, has the potential to predict production as much as learner choices. Statistical analyses such as regression analysis (as shown below) provide insights into the relative contribution of choices versus input features. Since the learners in this study select the input they receive, rather than examining input within an interactionist framework (e.g., recasts, clarification requests) that would be compatible with a social interactive research design, the present study examines the input's linguistic features (e.g., lexical density, structural complexity; Ellis & Schmidt, 1997) in providing an understanding of the complex interplay between choices, input, and learner output.

**Design Features Promoting Autonomy in CALL**

Schwienhorst (2003) outlines for CALL researchers and developers three approaches to promoting learner autonomy. In the individual-cognitive approach, reflective processes are aided through the act of writing (as opposed to speaking). In the social-interactive approach, interactions with peers or native speakers such as in project-based tasks promote autonomy. In the experimental-participatory approach, learners are made to be their own agents, and their own actions and choices promote autonomous learning. The

experiment presented below contains all three of these design features.

## Studying Autonomy and L2 Production in a 3D World Task

The present study on the relationship between autonomy in a 3D world task and L2 production incorporates the design principles outlined in the section above. And, while the experiment as a whole encourages learner autonomy in ways that are consistent with Schwienhorst (2003), the empirical study presented limits itself to studying the potential effects of learner autonomy on L2 production by focusing on learners' choices and those decisions' immediate consequences (i.e., the subsequent input they receive). First, a task-based 3D CALL environment encouraging learners to make choices about the L2 input they process aims at promoting an experimental-participatory approach. User-tracking technologies record learners' actions, allowing the researcher to study the choices they make (Fischer, 2007). Second, synchronous computer-mediated communication (SCMC) related to a 3D task provides an avenue for the promotion of the individual-cognitive approach to autonomy and the social interactive approach through written L2 expression while interacting with a peer to achieve a common goal. SCMC data provides a convenient, text-based format for examining learners' production in relationship to their choices in the 3D environment. The following section reviews the CALL literature, as it relates to autonomy, on user-tracking technologies, SCMC, tasks and 3D environments. The research questions and hypotheses follow.

### User-tracking technologies as a tool for studying learner autonomy

User-tracking technologies capture and archive users' actions or moves along with time and date stamps. In a 3D world program, user-tracking technologies can document various learner decisions: which NPCs they interact with, how long they read input that they encounter, and so forth. Tracking might also record the number of times a learner returns to retrieve input provided by an object or NPC. In short, tracking technologies can help us understand self-determination in a 3D learning environment, showing us the choices learners make to access L2 input that they themselves decide to process and (perhaps) contemplate. The present study uses these technologies to examine learners' choices about the L2 input to which they expose themselves.

CALL researchers have employed user-tracking technologies for a number of purposes, including whether learners fully take advantage of available software components and features (Fischer, 2007). From a learner autonomy perspective, user-tracking technologies can help us to measure the extent to which learners will decide to be 'agents' in CALL. Cobb and Stevens (1996) discovered that few students employed the optional software components of reading courseware even after practicing them and understanding their benefits. Hsu, Chapelle, and Thompson's (1993) learners neglected to use a component of an English grammar program that would aid with problematic answers. Pujolà (2002) found that learners seldom explored strategy-training components embedded in listening and reading comprehension courseware. However, word translators or dictionaries are frequently employed (Bland, Noblitt, Gay, & Armington, 1990; Davis & Lyman-Hager, 1997; Laufer & Hill, 2000; Noblitt & Bland, 1991), although learners may avoid this feature to complete activities more quickly (Hulstijn, 1993). Overall, despite the inclusion of software components that could aid learners in becoming more autonomous, the research employing user-tracking technologies shows that few learners utilize them consistently (Fischer, 2007).

### SCMC and Learner Autonomy

Schwienhorst (2003) argues that writing is one medium that facilitates reflective processes and self-critiques, and so tasks incorporating this mode of communication promote individual cognition. SCMC, such as iChat, provides such conditions. Research shows that SCMC engages learners in collaborative knowledge construction (Beauvois, 1997; Berge & Collins, 1995; Meunier, 1994; Warschauer, 1996, 1997). SCMC research is keenly focused on the nature of the L2 that dyads produce (Abrams, 2003; Blake, 2000; Collentine, 2009; de la Fuente, 2003; Keller-Lally, 2006; Smith 2003, 2005; Sotillo, 2000).

### *TB-SCMC in 3D environments*

SCMC can promote autonomy when dyads are charged with a common task, termed task-based SCMC (TB-SCMC; Collentine, 2009). Doughty and Long (2003) suggest that computer-mediated communication (CMC) activities should follow task-based language teaching (TBLT) design principles, in which "meaning is primary; there is a relationship to the real world; task completion has some priority; and the assessment of task performance is in terms of task outcome" (Skehan, 1996, p. 38). As far as autonomy is concerned, one potential advantage for learners engaged in TB-SCMC using a technology like iChat is enhanced awareness through L2 reflection (Schwienhorst, 2003): they can view their discourse during its production and refer to and even modify comments made at earlier points in the conversation because a running, written record is available to interlocutors.

Scant research is available on TB-SCMC set in 3D environments, although the little that exists has provided some insights into autonomy in CALL. Using synchronous and asynchronous communication in a virtual world, Zheng, Young, Brewer and Wagner (2009) investigated the effects of task completion on self-efficacy (i.e., learners' beliefs about their capabilities to complete a task; Bandura, 1977). They found that learners solving tasks within a virtual world reported higher levels of self-efficacy than learners in a control group. Three studies examine autonomy in relation to learners' usage of their avatar's communicative features (e.g., waving to get another avatar's attention), which demonstrate learners' propensity for taking control over their learning. Toyoda and Harrison (2002) as well as Peterson (2005) found that few learners employed their avatar's communicative features. Yet, Peterson's (2006) questionnaire revealed that, while most learners employed waving, few used other communicative features because the rapid exchanges in the chat impeded their use. More extensive use of user-tracking technologies could provide a better picture of learners' propensity for experimentation and self-determination in these learning contexts.

Interestingly, although SLA research ultimately attempts to understand how learners develop communicative competence in an L2, few researchers have explored either the relationship between autonomy and learner output or TB-SCMC in 3D environments and output. Sykes (2009) examined the effects of tasks in a 3D world on pragmatic development, namely, learner requests in Spanish; her analysis was unable to corroborate a facilitative effect for this environment on the pragmatic feature in question. Jones, Squires, and Hicks (2007) investigated the effects of a 3D environment on Japanese learners' pronunciation. Learners were to complete a series of questions in order to get a table at a restaurant (the task); the software recorded learners' responses and evaluated their approximation to standard Japanese. Beginning learners had greater gains in pronunciation than intermediate learners after using the 3D environment.

### Research Question and Hypothesis

This study contributes to the research on learner autonomy by studying the NPCs and objects with which learners choose to interact in TB-SCMC in a 3D world and the possible impact of such decisions on learner production. Advanced foreign language learners of Spanish participated in two tasks consisting of a 3D exploration segment followed by an SCMC chat segment (see Tasks below). Each 3D segment was a virtual environment providing opportunities to interact in an unsupervised fashion with various NPCs and objects for gathering information. Since participants were only compelled to interact with the 3D world by the demands of the task motivating them, interactions constitute choices—an important indicator of autonomous learning (cf., Dam 1995)—that each individual participant makes toward the task's goal. The study operationalizes choice by measuring with user-tracking technologies these self-determined interactions, referring to them as 'autonomous moves' (see Analysis below for a complete definition). As mentioned above, the input learners receive as a result of those choices could also mitigate their output. The study explores the relationship between autonomy and learner production by measuring the extent to which 'autonomous moves' and/or resulting input features predict linguistic complexity and accuracy in

the written output of the participants during the chat portion of the task.

Thus, this study asks the following:

In task-based CALL, do the type and quantity of autonomous moves and/or resulting input features while exploring a 3D environment predict linguistic complexity and/or accuracy in SCMC production?

The study employs four regression analyses to study the relationship between autonomy, input, and L2 production. The analyses' predictor variables (i.e., independent variables) are various calculations from the user-tracking data in the 3D segment; the response variables (i.e., dependent variables) are four measures of linguistic complexity and accuracy from the SCMC data. The researcher also provides a qualitative analysis of the data set.

## METHOD

### Participants

A total of 58 third-year university-level learners at a medium-sized university in the United States from intact classes participated in the study. They were enrolled in a junior-level course designed to review major grammatical structures and in a junior-level course focusing on developing conversational fluency. All participants had met or exceeded the learning outcomes from the previous course, a fourth-semester Spanish course. The classes were traditional, face-to-face (FTF) courses employing group activities as well as a variety of multimedia activities (e.g., watching videos, Internet exploration/research). While the courses required writing, speaking, and inductive/exploration activities, they did not entail any chat or instant messaging beyond the segments of this study. The researcher was the instructor of these classes but did not participate in the experimental tasks. The tasks were integrated into two lesson plans lasting two class periods of 1.5 hours for a total of three hours. No grades were awarded for participation since the tasks were incorporated into the class syllabus.

### Tasks

Learners participated in two tasks designed by the present author, each containing a 3D exploration segment and a subsequent SCMC segment. The 3D segments were authored in the Unity game development tool (http://unity3d.com/unity/). The SCMC segments occurred in a local area network via iChat, a synchronous conference application. The laboratory where both segments took place was equipped with individual Mac laptops placed on conference tables arranged in a semicircle.

Each task first took learners to a 3D island containing NPCs and objects relevant to a crime that was committed. Learners explored the island to collect clues as first-person characters (FPCs), queried NPCs, and collected clues from objects in the environment (e.g., notes, letters). The first task asked learners to find clues to solve a missing persons case, while the second task required learners to solve a murder-mystery. To explore the island, learners used the arrow keys for horizontal movement (e.g., walking) and the space bar for vertical movement (e.g., jumping). The 3D tasks contained two user interfaces within which to gather clues and information, one involving NPCs and another involving objects. Participants roamed within the 3D environment and freely chose what to approach and how often.

When learners approached an NPC, they were prompted with three possible questions that they could ask, by clicking on any of the three questions—written in Spanish—within a button, whereupon they received a written answer—again, in Spanish.[1]

*Figure 1.* Sample NPC question dialog.

When learners approached an object, they were presented with a written message in Spanish containing information to read (e.g., a learner could approach a diary to see an entry).
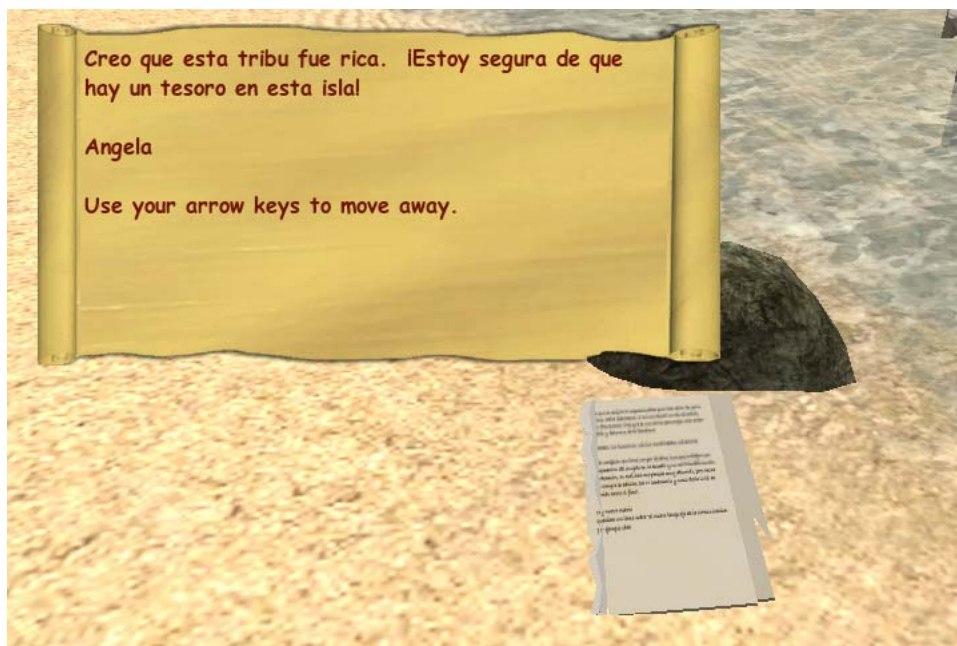


*Figure 2.* Sample message dialog.

The researcher emphasized that they should explore the environment all they wanted to solve the task. In both 3D segments, the researcher directed the learners to close the application after 10 minutes of gathering clues.

Regarding the SCMC segments, after each task students were paired up into predetermined, randomly assigned dyads. Each dyad was to come to a consensus relating to the relevant crime: for the missing-persons task, dyads were to determine the reason(s) for the person's disappearance; for the murder-mystery, dyads had to determine the reason(s) for the murder. The chat phase lasted 25 minutes.

To familiarize the learners with both the 3D technology and iChat, the day before the experiment the learners navigated a sample 3D world (not employed in the present analyses) containing examples/instances of the technologies described here. Subsequently, the participants practiced chatting about the sample 3D world.

The data for assessing linguistic complexity and accuracy took place in an SCMC environment because it offers processing conditions favorable to learners generating linguistic complexity. Resource depleting processing, such as FTF conversations, works against the production of linguistic complexity (Skehan, 1996). SCMC requires a certain degree of spontaneity at a pace similar to FTF interactions (Smith, 2003). Yet, SCMC lessens the pressure that unfilled pauses place on the need for spontaneity in FTF oral interactions, which can lead to incomplete processing of complex linguistic elements due to the limited resources available to short-term memory (Kern, 1995; Warschauer & Kern, 2000). Sotillo (2000) as well as Keller-Lally (2006) found that the delayed nature of asynchronous discussions led to the production of more syntactically complex language. There is also evidence that participants' production may contain more complex language in SCMC than in oral, FTF tasks (Kern, 1995; Salaberry, 2000).

**Dataset: User-Tracking Data and Chatscripts**

Two types of user-tracking data were collected for the study. Each 3D application tracked the learners' movements, actions, and choices to a MySQL database with Javascript and PHP. The applications also tracked the time at which each recorded event occurred, permitting the calculation of the number of types of events a learner's movements or choices affected as well as the amount of time events took and how often they were repeated. The database represented a total of 7,339 recorded events.

Each dyad's iChat transcript was archived to a text document and collated by participant, allowing the measurement of various aspects of a learner's linguistic complexity and accuracy. The corpus totaled 27,315 words.

**Analysis**

A mixed-method analysis was employed for data analysis, combining quantitative and qualitative perspectives (Green & Caracelli, 1997). The quantitative analysis examined the extent to which features of the user-tracking data from the 3D segments predicted the production of four measures of linguistic complexity and accuracy in SCMC segments.

There were three types of predictor variables: (a) interaction-based features, (b) input features, and (c) time spent without interacting with NPCs and objects.

Table 1 represents measurements of learner interactions with NPCs and objects in the virtual environment. These variables measure various choices that the user-tracking technologies recorded. For instance, if a learner approached many NPCs, which would in turn present him or her with a series of question dialogs as in Figure 1, his or her *question events* count would be relatively high. If that same learner were to approach the question dialog in Figure 1 two or more times, his or her *repeated question events* count would increase. And, if s/he were to approach a message dialog such as that of Figure 2 more than once, his or her *repeated message events* count would increase.

Table 1. *Predictor variables: Interaction-based features affected by learners in 3D environment*

| Label | Description |
|---|---|
| Question events | The number of times a learner cued question dialogs (i.e., by approaching a NPC) or clicked a written question button to cue a written answer. |
| Question time | Time (in seconds) within a question dialog (i.e., where the screen showed question buttons to click), as well as when time was spent reading question buttons and their corresponding answers. |
| Message events | The number of times a learner cued message dialogs (i.e., by approaching an object) to receive written information (e.g., a note, letter). |
| Message time | Time (in seconds) spent reading the contents of message dialogs. |
| Total interaction events | The sum of question and message events. |
| Total interaction time | The sum of question and message time. |
| Average interaction time | Average time per question and message event. |
| Repeated question events | The number of times a learner cued a question event two or more times. |
| Repeated message events | The number of times a learner cued a message two or more times. |
| Repeated interaction events | The sum of repeated question events and repeated message events. |

The predictor variables in Table 2 represent calculations of the linguistic features of the questions, answers, and messages a student read in the virtual environment.

Table 2. *Predictor variables: Input features of 3D environment*

| Label | Description |
|---|---|
| Clauses read | Number of clauses read in questions, answers and messages. |
| Tokens read | Number of tokens (i.e., words) read in questions, answers and messages. |
| Reading type-token ratio | Type-token ratio of all questions, answers and messages read: the ratio of unique words to total words in input read. The higher the ratio, the more unique words the input had. |
| Reading lexical density | Lexical density of all questions, answers and messages read: the ratio of unique main parts-of-speech words (i.e., nouns, verbs, adjectives, and adverbs) to total main parts-of-speech words in input read. The higher the ratio, the more semantically dense the input was. |

Based on the recordings of the user-tracking technologies, the researcher was able to identify particular statements the participants read and assess the statements' linguistic characteristics. For instance, if a learner approached the question dialog in Figure 1, on that screen alone, s/he would read 4 clauses (i.e., each visible statement here has 1 clause) and 13 tokens (i.e., the statement *Vi a la mujer el jueves, ayer* plus the tokens in the visible buttons). The *type-token ratio* of the message in Figure 2 would be 0.88 (15 unique words/17 total words); calculating the ratio of total unique words to total words that a learner read

in all questions and dialogs would yield the learner's *reading type-token ratio*.

Finally, the variable titled *non-interaction time* represents the time (in seconds) when neither a question nor message dialog was cued on the screen, such as moving from one NPC to an object or simply viewing the 3D surroundings.

Three response variables were calculated from the chatscripts to measure linguistic complexity and one response variable to measure accuracy, all common to task-based research. The first two response variables represent measures of lexical complexity (cf., Ellis, 2003). The third variable represents a measure of structural complexity, and the fourth linguistic accuracy.

Table 3. *Response variables: Measures of linguistic complexity in learner production*

| Label | Description |
|---|---|
| Learner type token ratio | Type-token ratio of a learner's production: the ratio of unique words to total words produced in the SCMC segment. The higher the ratio, the more unique words a learner produced. |
| Learner lexical density ratio | Lexical density of a learner's production: the ratio of unique main parts-of-speech (i.e., nouns, verbs, adjectives, and adverbs) words to total main parts-of-speech words in input read. The higher the ratio, the more semantically dense the learner's production was. |
| Learner clauses per c-unit | A c-unit is an utterance containing a single complete sentence, phrase, or word and that has a clear semantic and pragmatic meaning in the context in which it occurs. This is similar to the T-unit, although it is more appropriate for the elliptical nature of conversations and SCMC (cf., Skehan, 1996). |
| Learner percentage of error-free clauses | The percentage of clauses a learner produced that contained no grammatical or lexical errors. All errors in syntax, morphology, and lexical choice were considered (cf., Ellis, 2003). |

The type token and lexical density ratios were calculated for each learner with concordance software. For example, a particular participant's type-token ratio was calculated by dividing the unique words s/he produced by the total number of words s/he produced in the SCMC portion of the experiment. The error-free clauses were derived for each learner by counting the frequency of clauses containing no errors divided by the total number of clauses the learner produced, which was calculated by totaling the number of independent and dependent clauses per learner. For example, if a learner produced a statement such as *creo que Angela *está *el culpable* ("I think Angela is the guilty one"), s/he would be counted as having one error-free clause (i.e., *creo que*) and one erred clause (i.e., *Angela *está *el culpable*). An independent clause represented the first clause of a c-unit. A dependent clause was headed by either a subordinating conjunction such as *que* ("that") or a coordinate conjunction such as *y* ("and") or *o* ("or").

An inter-rater reliability analysis was employed to check the construct validity of the researcher's coding of erred clauses and c-units since identifying both constructs requires a certain degree of judgment. Specifically, a fellow researcher with experience in this field was presented with a random sample of 100 segments from the corpus. She was to both count the number of c-units and the errorless clauses. Her two sets of scores were compared with those of the researcher with a Pearson correlation, as each of the 4 datasets ($n = 100$) was on an interval scale. Concerning the count of errorless clauses, the correlation between the two researchers was significant, $r = .92$, $p = .01$. Regarding the count of c-units, the correlation between the two researchers was also significant, $r = .91$, $p = .01$.

For all-subsets regression analyses were conducted, each employing one of the response variables and all

of the predictor variables. The process involves screening a set of potential predictor variables for standard assumptions of linear regression and then submitting the reduced set to an all-subsets analysis—rather than a stepwise procedure—to identify the so-called "best subset". Social scientists frequently employ stepwise procedures for building regression models; yet, many statisticians do not recommend stepwise analyses when the number of predictor variables is more than five or so (Rencher, 2002).

The researcher employed a standard data-screening process, identifying with scatter plots the potential predictor variables that had discernible correlations with the response variables. Three predictor variables had no correlation with any of the response variables, and so they were excluded from the all-subsets analyses: repeated question events, repeated message events, and non-interaction time.

Since the all-subsets analysis is computationally intensive and not available in many commercial software packages for the social sciences, the researcher used the R statistical package and its all-subsets regression package to identify the optimal predictor variables for each of the four response variables (see Dalgaard, 2002). SPSS then produced the regression statistics for the four resulting models.

Regarding the qualitative analysis, the researcher includes a narrative of two learners' interactions in the 3D environment and samples of the Spanish that they produced in the SCMC segments to provide a complementary depiction of the role of autonomy and learner production.

## RESULTS

Regarding the quantitative analysis,[2] the all-subsets analysis revealed that four of the 11 potential predictor variables (14 original variables less the three discarded by scatter-plot analyses) significantly predicted *learner type-token ratio*.

Table 4. *Regression analysis for response variable learner type-token ratio*

| Variable | Model *B* | 95% CI |
|---|---|---|
| Constant | 0.547** | [0.493, 0.601] |
| Message time | -0.001* | [-0.002, 0.001] |
| Clauses read | -0.001** | [-0.002, 0.000] |
| Reading lexical density | 0.108* | [0.001, 0.216] |
| Repeated interaction events | 0.013** | [0.004, 0.022] |
| $R^2$ | .085 | |
| *F* | 3.680** | |

*Note. *p < .05. **p < .01.*

Two predictor variables were interaction-based features affected by learners in the 3D environment. Message time (beta = -0.001), with its significant yet negative coefficient, was disassociated with the response variable. Repeated interaction events (beta = 0.013), with its significant, positive coefficient, was associated with the response. Two of the predictor variables were input features of the 3D environment. Clauses read (beta = -.001) was disassociated with the response variable, while reading lexical density (beta = .108) was associated with the response variable.

All told, regarding the relationship between autonomy and learners' lexical production, choosing to re-interview an NPC or re-read a message correlates with higher levels of lexical complexity in learner production, although this observation is most applicable to learners who spent little time reading those messages. Concerning input features, input with higher levels of lexical density was associated with more

lexical complexity in the learners' production, especially when learners read relatively few clauses overall.

The all-subsets analysis revealed that the same four predictor variables significantly predicted *learner lexical density ratio*.

Table 5. *Regression analysis for response variable learner lexical density ratio*

| Variable | Model *B* | 95% CI |
|---|---|---|
| Constant | 0.642** | [0.591, 0.693] |
| Message time | -0.001 | [-0.001, 0.000] |
| Clauses read | -0.001* | [-0.002, 0.000] |
| Reading lexical density | 0.098 | [-0.004, 0.200] |
| Repeated interaction events | 0.010* | [0.002, 0.019] |
| $R^2$ | .052 | |
| $F$ | 2.578* | |

*Note. *p < .05. **p < .01.*

Furthermore, the associations and disassociations of the four identified predictor variables were the same as in the learner type-token ratio analysis. While this may appear to be a redundant analysis, it nonetheless provides a certain level of corroboration for the above assessments of the relationship between autonomous moves and input features with respect to learners' lexical complexity.

The all-subsets analysis revealed that only two of the potential predictor variables significantly predicted *learner clauses per c-unit*.

Table 6. *Regression analysis for response variable learner clauses per c-unit*

| Variable | Model *B* | 95% CI |
|---|---|---|
| Constant | 27.087** | [21.347, 32.827] |
| Question time | 0.029* | [0.006, 0.053] |
| Reading type-token ratio | -14.587* | [-25.700, -3.474] |
| $R^2$ | .047 | |
| $F$ | 3.842* | |

*Note. *p < .05. **p < .01.*

One variable, question time (beta = .029), was an interaction-based feature, indicating that the more time learners spent interacting with the question dialogs (e.g., selecting and reading questions, reading their answers), the more structurally complex their output was. However, the reading type-token ratio (beta = -14.587) was disassociated with the production of structural complexity. Thus, the positive effect of interacting with question dialogs on the structural complexity of output occurs when the lexical variety of what learners read is low.

The all-subsets analysis revealed that four of the potential predictor variables significantly predicted *learner percentage of error-free clauses*.

Table 7. *Regression analysis for response variable learner percentage of error-free clauses*

| Variable | Model $B$ | 95% CI |
|---|---|---|
| Constant | 1.924 | [1.619, 2.228] |
| Message events | -0.025* | [-0.049, -0.002] |
| Message time | 0.003* | [0.000, 0.007] |
| Clauses read | 0.005* | [0.001, 0.009] |
| Reading type-token ratio | -0.865** | [-1.496, -0.233] |
| $R^2$ | .047 | |
| $F$ | 2.422* | |

*Note. *p < .05. **p < .01.*

As in the first two regression analyses above, two of the predictor variables were interaction-based features affected by learners in the 3D environment while the other two constituted input features. Concerning the interaction features, message events (beta = -0.025) were disassociated with the response variable, while message time (beta = 0.003) was correlated positively. Regarding the two input variables, clauses read (beta = 0.005) was associated with the predictor variable while reading type-token ratio (beta = -0.865) was disassociated.

All told, this is the only model that hints that interacting with messages—specifically, the amount of time reading messages—facilitates production in the sense that it is associated here with more error-free production. However, as in the two lexical-complexity analyses, this model also indicates that, concerning accuracy, there is little benefit to reading numerous message dialogs. This analysis finds a facilitative effect for clauses read—as opposed to the two analyses on learners' lexical complexity—on accuracy. In contrast to the two analyses on the predictors of the learners' lexical complexity where this was associated with greater lexical complexity in the input, lexical complexity in the input in this analysis was associated with more errors, as was the case in the learner clauses per c-unit analysis. Thus, lexical complexity in learner input is associated with more lexical complexity in learner output, but lexical complexity in input is not associated with structural complexity or accuracy.

Regarding the qualitative analysis, a perusal of various learners' 3D and chat activities revealed that not all learners explored the 3D world alike. The following describes two learners, DT and JK, whose choices and production differ. Still, it is important to note that there were no significant differences in terms of each learner's total amount of time in the 3D segment: DT spent 678 seconds (11.3 minutes) in the 3D world while JK spent 718 seconds, or 12 minutes, $\chi^2(1) = 1.15$, $p = .28$. Yet, DT had a total of 114 interaction-based events while JK logged a total of 79, a difference that was significant, $\chi^2(1) = 6.35$, $p = .01$. DT interacted with significantly more question events (49 total) than message events (20 total), $\chi^2(1) = 12.19$, $p = .01$, while there were no significant differences between the number of question events (27 total) and object events (19 total) for JK, $\chi^2(1) = 1.39$, $p = .24$.

Specifically, DT's first move entailed approaching an NPC and reading the three possible questions and answers, spending on average 4.7 seconds with each question event. He chose to approach another NPC and did the same, spending 5.8 seconds on average per question event. Next, he interacted with a message dialog, moved away, and then returned to it immediately afterwards, spending a total of 8.5 seconds on this message. After 2.7 seconds, DT interviewed another NPC for a total of 33 seconds, viewing all three of the dialog's question-answer pairs, and one pair twice. A mere 1.6 seconds later, DT approached an object and read its message twice (i.e., entering its vicinity and cuing its dialog twice) for a total of 6.5 seconds; then he proceeded to interview another NPC, and then another NPC, and another object,

seemingly to explore all of the possible interactions during the 10-minute 3D segment. DT's interactions in the 3D world remained the same throughout the activity, and his choices led him to interactions with NPCs and objects that provided him with the input necessary to solve the task. He did not miss any clues and interviewed all of the NPCs rather quickly, returning to a few of the NPCs and objects near the end of the 3D activity.

In his chat, DT provided lexically and syntactically complex hypotheses with few errors, stating, *Sí, pienso que era Tito también. Pero, ?por qué? ?Qué es la relación?* "Yes, I think it was Tito too. But, why? What is the relationship?". DT then reasons: *pero nadie le cae bien a Pedro* "but Pedro doesn't like anyone", employing the complex gustar-like construction. DT is not quoting what people said; everything he writes is in the abstract.

JK spent almost the same amount of time interacting with question events (483 seconds) as did DT (469 seconds), $\chi^2(1) = 0.21$, $p = .65$; however, DT had 49 question events and JK logged significantly fewer, 27 events, $\chi^2(1) = 6.37$, $p = .01$. JK interviewed only four of the six NPCs and would, with a few exceptions, spend quite a bit of time reading each question-answer dialog. In one instance, JK spent 91.4 seconds (about 1.5 minutes of the 10 minutes spent on the task) reading the question-answer dialogue for one NPC; in another she spent 79.2 seconds. She approached two NPCs three consecutive times each (i.e., moving in and out of their respective vicinities to re-prompt their respective dialogs), but only apparently read the questions; she never cued any of the answers. DT never exhibited this move: every time he approached an NPC, he clicked on a question to see its answer.

JK also interacted differently with message events (i.e., objects). For example, she spent 136.3 seconds (2.23 minutes) with just three message events, apparently reading each an average of eight times. In one instance, she interacted with the same message (i.e., *Esta es la ropa que se puso Nora para bailar en la fiesta.* "These are the clothes that Nora put on for the dance at the party") nine consecutive times for a total of 45.8 seconds.

JK's production contains little lexical or syntactic complexity, various errors, and code-switching, such as: *pero quién 'would' querer 'kill' su compañero o amigo?* "But who would want to kill his roommate or friend?" As another example: *no. yo no creo. Tito dijo que el bailé con Nora para los 30 minutos.* "No. I don't think so. Tito said that I danced (should be: he danced) with Nora for the 30 minutes."

Clearly, there were differences in these learners' proficiency level, and it remains to be researched whether L2 proficiency affects the manner and amount learners interact with L2 materials autonomously. Nonetheless, this admittedly cursory quantitative analysis suggests that learners who do not actively seek out task-relevant input are less prepared to communicate fully about the task, even within an SCMC environment offering more time to process individual utterances.

**DISCUSSION AND CONCLUSIONS**

Since it has been posited that autonomy positively affects language learning (Schwienhorst, 2002, 2003, 2008), this study examined whether learners' choices (Dam, 1995) in a task-based, 3D environment affect learner production in terms of complexity and accuracy. Third-year university learners of Spanish interacted with NPCs and objects to gather information and then chat with a classmate in an SCMC environment. The tasks required dyads to solve a murder mystery and a missing-persons case. Moving within the 3D environment, the task compelled learners to interview NPCs and obtain information from objects, gathering written input from both. The amount of interaction with NPCs and objects was a function of self determination—some learners chose to gather more data and return to data more often than others.

Regarding interaction variables, learners exhibiting high levels of exploratory strategies by re-interviewing an NPC or reflective strategies by re-reading a message in the 3D segment yielded relatively

high levels of lexical complexity. Time spent reflecting on written input has varying effects on output. On the one hand, spending time interacting with question-answer dialogs is associated with syntactic complexity. On the other hand, large amounts of time reading messages do not seem to help lexical complexity. The data are not clear-cut with respect to accuracy; interacting with messages seems to affect the amount of error-free clauses learners produce if learners do not read a large number of messages.

Nonetheless, the characteristics of the input learners receive as a result of their choices also influences learner production. The lexical density of the input influences the lexical complexity of learner production if the input does not contain numerous clauses. Reading numerous clauses seems to facilitate learners' abilities to generate error-free clauses, but only if learners are reading sentences with little lexical variety (i.e., low type-token ratios). However, having a low lexical variety can have a positive impact on syntactic complexity because the analysis shows that reading input with lower type-token ratios is associated with the production of more clauses per c-unit.

All told then, the statistical analysis indicates that self determination in combination with the nature of the input learners receive determines the complexity and accuracy of what they produce. The analysis reveals a complex interaction between autonomy, input, and production.
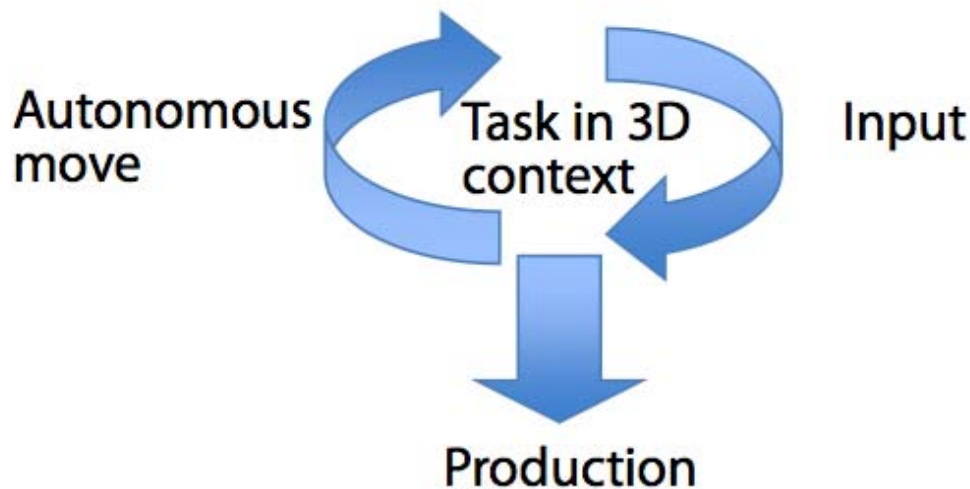


*Figure 3.* Autonomy, task, input and production.

Making more choices in a 3D environment does not necessarily lead to the production of linguistic complexity or accuracy. Both the learners' choices and the subsequent input they receive affect their production. This is not a trivial observation, since it implies that CALL developers should not only provide opportunities for learner autonomy, they should also consider the features of and access to input learners will have. 3D worlds providing a visually and culturally authentic experience for learners (e.g., a garden centrally located in a home) may have less effect on production than the extent to which such environments encourage autonomous learning and the type of input they receive in the 3D environment. Additionally, learners may need a well-defined task to complete in the virtual environment to encourage autonomy.

Despite the study's larger sample size, there are limitations to its generalizability. While all learners were

enrolled in third-year courses, the variability in the participants' overall proficiency in Spanish is not accounted for. In future studies a proficiency level predictor variable in the analysis could reveal whether proficiency level in speaking and/or reading mitigates autonomy, input processing, and production. That said, it is important to keep in mind that the correlations between individual predictor variables and individual response variables were highly significant. Additionally, as is common to TBLT research, the measures representing linguistic complexity and accuracy are general in nature, and surely mask more subtle effects of autonomy and input on production. Future research could explore in greater depth this relationship, and even consider the factors of comprehension and intake. Technology that records learners' moves and tracks what they read (and perhaps even hear) can help us to understand the effects of virtual input not only on production but also on comprehension abilities. Finally, regression analysis is considered a first step in establishing causality (Tabachnick & Fidell, 2001). A study employing this data set with structural equation modeling or an independent sample verifying the study's regression models could test the strength of the causality relationships explored in the conclusions.

## NOTES

1. Technically speaking, learners navigated the virtual environment with keyboard controls. NPCs and objects (e.g., diary entries) were surrounded by transparent collider boxes, which triggered the appearance of controls such as buttons (e.g., questions whose answer appeared once clicked) and non-modifiable textboxes (e.g., written answers to questions, written information).

2. Three of the models produced by the all-subsets analysis contained four predictor variables, and one contained two predictor variables. Given the sample size of 58 and the number of cases in sample per variable of 116, the three models with four predictor variables averaged 29 cases per predictor variable and the two variable model averaged 58 cases per predictor variable. These cases-per-predictor ratios represent adequate sample sizes to detect reliable correlation estimates (Tabachnick & Fidell, 2001). Additionally, none of the four data sets contained multicollinearity (or singularity) violations, such that no two variables correlated at $r = 0.90$ or higher.

## ABOUT THE AUTHOR

Dr. Karina Collentine is Assistant Professor of Spanish and Secondary Spanish Education at Northern Arizona University.

E-mail: Karina.Collentine@nau.edu

## REFERENCES

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

Bandura, A. (1977). *Social learning theory*. New York: General Learning Press.

Beauvois, M. (1997). High tech, high touch: From discussion to composition in the networked classroom. *Computer Assisted Language Learning, 10,* 57–69.

Berge, Z., & Collins, M. (Eds.), (1995). *Computer mediated communication and the on-line classroom in distance learning*. Cresskill, NJ: Hampton Press.

Blake, R. (2000). Computer-mediated communication: A window on L2 Spanish interlanguage.

*Language Learning & Technology*, *4,* 120–136. Retrieved from http://llt.msu.edu/vol4num1/blake/default.html

Bland, S., Noblitt, J., Gay, G., & Armington, S. (1990). The naïve lexical hypothesis: Evidence from CALL. *Modern Language Journal, 74,* 440–450.

Cobb, T., & Stevens, V. (1996). A principled consideration of computers and reading in a second language. In M. Pennington (Ed.), *The power of CALL* (pp. 115–136). Houston, TX: Athelstan.

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

Dalgaard, P. (2002). *Introductory statistics with R*. New York: Springer.

Dam, L. (1995). *Learner autonomy* 3: *From theory to classroom practice,* Dublin, Ireland: Authentik.

Darasawang, P., & Reinders, H. (2010). Encouraging autonomy with an online language support system. *CALL-EJ Online, 11,* Retrieved August 29, 2010, from http://callej.org/journal/11-2/darasawang_reinders.html.

Davis, J., & Lyman-Hager, M. (1997). Computers and L2 reading: Student performance, student attitudes. *Foreign Language Annals, 30,* 58–72.

de la Fuente, M. (2003). Is SLA interactionist theory relevant to CALL? A study on the effects of computer-mediated interaction in L2 vocabulary acquisition. *Computer Assisted Language Learning, 16,* 47–81.

Doughty, C., & Long, M. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology, 7*, 50–80. Retrieved from http://llt.msu.edu/vol7num3/doughty/default.html

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

Ellis, N., & Schmidt, R. (1997). Morphology and longer-distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition, 19,* 145–172.

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behavior in *CALL. Computer Assisted Language Learning, 20,* 409–442.

Greene, J. & Caracelli, V. (1997). Defining and describing the paradigm issue in mixed-method evaluation. In J. Greene & V. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms* (pp. 5–18). New Directions for Program Evaluation, No. 74. San Francisco, CA: Jossey-Bass.

Gick, C. (2002). Autonomes Grammatiktraining in der Mediathek. *Babylonia*, *2*, 55–58.

Hsu, J., Chapelle, C., & Thompson, A. (1993). Exploratory learning environments: What are they and do students explore? *Journal of Educational Computing Research, 9*, 1–15.

Hulstijn, J. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *Modern Language Journal, 77,* 139–147.

Jones, G., Squires, T., & Hicks, J. (2007). Combining speech recognition natural language processing with 3D online learning environments to create distributed authentic and situated spoken language learning. *Journal of Educational Technology Systems*, *39*, 375–392.

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. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. *Modern Language Journal, 79,* 457–476.

Little, D. (1991). *Learner autonomy 1: Definitions, issues, and problems*. Dublin, Ireland: Authentik.

Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language Learning & Technology, 3*, 58–76. Retrieved from http://llt.msu.edu/vol3num2/laufer-hill/index.html

Meunier, L. (1994). Computer-assisted language instruction in cooperative learning. *Applied Language Learning*, *5,* 31–56.

Noblitt, J., & Bland, S. (1991). Tracking the learner in computer-aided language learning. In B. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 120–131). Lexington, MA: D. C. Heath.

Pantelidis, V. (1993). Virtual reality in the classroom. *Educational Technology, 33*, 23–27.

Peterson, M. (2005). Learning interaction in an avatar-based virtual environment: A preliminary study. *PacCALL Journal, 1,* 29–40.

Peterson, M. (2006). Learner interaction management in an avatar and chat-based virtual world. *Computer Assisted Language Learning, 19,* 79–103.

Pujolà, J. T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL, 14,* 235–262.

Reinders, H. (2006). Supporting self-directed learning through an electronic learning environment. In T. Lamb & H. Reinders (Eds.), *Supporting independent learning: issues and interventions* (pp. 219–238). Frankfurt am Main: Peter Lang Publishing.

Reinders, H. (2007). Big brother is helping you: Supporting self-access language learning with a student monitoring system. *System, 35,* 93–111.

Rencher, A. (2002). *Methods of multivariate analysis*. New York: Wiley-Interscience.

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

Schwienhorst, K. (2002). Why virtual, why environments? Implementing virtual reality concepts in CALL. *Simulation and Gaming*, *33,* 196–209.

Schwienhorst, K. (2003). Learner autonomy and tandem learning: Putting principles into practice in synchronous and asynchronous telecommunications environments. *Computer Assisted Language Learning, 16*, 427–443.

Schwienhorst, K. (2008). *Learner autonomy and CALL environments*. New York: Routledge.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, *17,* 38–62.

Smith, B. (2003). Computer-mediated negotiated interaction: An expanded model. *Modern Language Journal, 87,* 38–57.

Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly, 39,* 33–58.

Sotillo, S. (2000). Discourse functions and syntactic complexity in synchronous and asynchronous

communication. *Language Learning & Technology, 4,* 82–119. Retrieved from http://llt.msu.edu/vol4num1/sotillo/default.html

Sykes, J. (2009). Learner requests in Spanish: Examining the potential of multiuser virtual environments for L2 pragmatics acquisition. In L. Lomicka & G. Lord (Eds.), *The next generation: Social networking and online collaboration* (pp. 199–234). San Marcos, TX: Texas State University: CALICO.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Needham Heights, MA: Allyn and Bacon.

Toogood, S. & Pemberton, R. (2002). Integrating self-directed learning into curriculum: A case study. In P. Benson & S. Toogood (Eds.), *Challenges to research and practice* (pp. 86–110). Dublin: Authentik.

Toyoda, E. & Harrison, R. (2002). Categorization of text chat communication between learners and native speakers of Japanese. *Language Learning & Technology*, *6*, 82–99. Retrieved from http://llt.msu.edu/vol6num1/toyoda/default.html

van Lier, L. (2010). Forward: Agency, identity, and language learning. In B. O'Rourke & L. Carson (Eds.), *Language learner autonomy: Policy, curriculum, classroom. A Festschrift in honor of David Little* (pp. ix–xiv). Oxford: Peter Lang.

Warschauer, M. (1996). Motivational aspects of using computers for writing and communication. In M. Warschauer (Ed.), *Telecollaboration in foreign language learning: Proceedings of the Hawai'i Symposium* (29–46). Honolulu, HI: University of Hawai'i, Second Language Teaching & Curriculum Center.

Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *Modern Language Journal, 81,* 470–481.

Warschauer, M., & Kern, R. (Eds.). (2000). *Network-based language teaching: Concepts and practice*. Cambridge: Cambridge University Press.

Zheng, D., Young, M., Brewer, R., & Wagner., M. (2009). Attitude and self-efficacy change: English language learning in virtual worlds. *CALICO Journal, 27*, 205–231.