ReCALL (2020), 32: 3, 232–249
doi:10.1017/S0958344020000026

ARTICLE

The effects of pre-task planning on EFL learners' oral performance in a 3D multi-user virtual environment

Julian ChengChiang Chen
Curtin University, Australia (Julian.Chen@curtin.edu.au)

Abstract

Prior research on pre-task planning examines its effects on the quality of second language (L2) learners' planned output. Planning mitigates the cognitive overload placed upon L2 learners' oral performance, thus improving language production. Despite the pedagogical benefits, studies on pre-task planning on L2 learners' oral output are conducted mostly in a lab or class setting. Whether or not similar effects of pre-task planning can be evidenced in three-dimensional (3D) multi-user virtual environments (MUVEs), such as Second Life (SL), is still less explored. Hence, this study investigates whether pre-task planning could enhance the quality and quantity of English as a foreign language (EFL) learners' task-oriented, voice-based outcomes in SL. Nine EFL learners worldwide participated in this 10-session virtual class. Data were collected through students' oral presentations in performing real-life simulated tasks related to their home cultures and interests. Yuan and Ellis's (2003) framework of T-units measures was adopted to analyze their linguistic performance measured by complexity and accuracy. Results indicated that EFL learners showed statistically significant improvement on grammatical complexity on the levels of syntactic complexity and variety (but not on lexical variety) and on linguistic accuracy across all measured levels (error-free clauses/T-units/verb forms). It is suggested that pre-task planning can be seeded in task-based instruction either in a classroom-based or 3D MUVE setting to optimize the quality of learners' linguistic performance. Tasks that are real-world oriented and targeting learners' cultural repertoires and world knowledge also positively impact their virtual learning experiences. These significant implications add new research and pedagogical dimensions to the field of computer-assisted language learning.

Keywords: 3D multi-user virtual environment (MUVE); Second Life (SL); pre-task planning; English as a foreign language (EFL); task-based language teaching (TBLT)

1. Introduction

In today's digital era, language learners are wired with all forms of technology anytime and anywhere. Our digital generation invests substantial time online, frequents social networking spheres, plays online games, and multitasks with assignments while emailing and texting (Prensky, 2005a, 2005b). Three-dimensional (3D) multi-user virtual environments (MUVEs), such as World of Warcraft, Active Worlds, and SimCity, have gained popularity among the Net Generation due to their affordances augmented by simulation, immersion, creativity, and collaboration (Peterson, 2016a; Puentedura, 2006; Sadler, 2012). Second Life (SL) has also drawn the attention of second language (L2) learners to explore this vibrant 3D space and interact with other users in world languages. It offers a pedagogical avenue to realize real-life tasks, from dining in a 3D Italian restaurant to taking a virtual field trip to Machu Picchu. Designing technology-enhanced, task-based instruction in 3D MUVEs has enabled educators to experiment with

Cite this article: Chen, J.C.C. (2020). The effects of pre-task planning on EFL learners' oral performance in a 3D multi-user virtual environment. *ReCALL* 32(3): 232–249. https://doi.org/10.1017/S0958344020000026

innovative ways of teaching, and to stimulate learners to practice target languages beyond the walls of their class. In addition, it promotes the use of language for communicative, meaningful, and experiential purposes (Chen, 2016a, 2016b, 2018; Chun, Kern & Smith, 2016; Dawley & Dede, 2014; Sadler & Dooly, 2013).

Prior research on pre-task planning explored the role played in L2 learners' interlanguage development and the quality of their planned output (Crookes, 1989; Ellis, 1987; Ortega, 1999). The effects of pre-task planning are reported to mitigate the cognitive overload of the task demands for L2 learners on their speech outcomes measured by complexity, accuracy, and fluency (Ellis, 2009a; Foster & Skehan, 1996; Skehan, 1996; Yuan & Ellis, 2003). The opportunity of time provided for L2 learners to plan before coming to terms with the higher levels of task demands has theoretical and pedagogical implications in the field of second language acquisition (SLA). This has not only illuminated how learners' internal planning processes can benefit their noticing of L2 forms and monitoring the output but also motivated language teachers to incorporate the concept of pre-task planning in instruction to optimize L2 learners' interlanguage development and usage. Despite the positive claims, previous studies on pre-task planning were mostly conducted in a physical class or a lab setting (Mehnert, 1998; Skehan & Foster; 2005; Tajima, 2003). Less research attention has been directed toward the effects of pre-task planning on English as a foreign language (EFL) learners' oral output in 3D MUVEs, much less targeting EFL learners with culturally/linguistically diverse backgrounds. This under-researched area serves as a catalyst for this study, rendering a new research avenue to investigate whether pre-task planning could be a facilitative factor in enhancing the quality and quantity of EFL learners' task-oriented, voice-based outcomes in SL.

2. Research background

2.1 Task-based design in Second Life

SL, developed by Linden Lab in 2003, is a 3D MUVE that allows its users (used interchangeably here with residents) to utilize self-created avatars to interact with other users. According to the analytics report released by Linden Lab (2013), SL has drawn more than 36 million users worldwide to this 3D virtual sphere since its launch, and the number is still increasing. Depending on one's personal preferences and creativity, SL residents can change their avatar's appearances and outfits to make their identities versatile and surreal. For example, they can instantaneously change their avatar's representation into a superhero, an animal, or a person that may or may not resemble their true self in real life. Avatars in SL can teleport to various virtual islands (in land) through SLURLs (SL teleport links) to designated SL locations with just the click of a mouse. SL also affords residents to walk, fly, and socialize with others via public text chat, voice chat, private instant message, or performing paralinguistic cues, such as laughing and dancing.

Another unique feature of SL is that the residents can build 3D objects (e.g. house, clothing) using scripting functions and take snapshots of in-land activities on the fly (Sadler & Dooly, 2013). Similar to the real world, SL enables avatars to participate in a variety of social events that simulate real-life routines (e.g. attending a conference or visiting Times Square) (Wang & Burton, 2013). Due to viable affordances in 3D simulation, immersion, tele/copresence, and multimodality, language learners can easily access SL without the constraints of time and physical boundaries (Canto, de Graaff & Jauregi, 2014; Cooke-Plagwitz, 2009; Peterson, 2016a). Specifically, it allows them to use a target language to simultaneously and spontaneously interact with speakers from culturally/linguistically diverse backgrounds as if they were in the real world (Lee & Gerber, 2013; Peterson, 2012, 2016b). Hence, the flexibility, low cost, and vibrant features afforded by SL have attracted a growing number of language learners to create their own second life with hopes of practicing their target language for communicative and authentic purposes.

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

The methodological principles of task-based language teaching (TBLT), as argued by Doughty and Long (2003), are theoretically sound (e.g. focus on meaning and form [not on forms], rich input, and authentic tasks) and pedagogically driven (e.g. problem-solving, collaborative learning, and individualized instruction), with great potential for online language instruction. Learning by doing, one of the integral TBLT principles, is well suited for the immersive nature of SL, which augments reality and deepens one's learning experience (Puentedura, 2006). Interactive tasks also draw learners' attention to linguistic forms that need refinement, leading to better quality in language production as measured by accuracy and complexity (Swain & Lapkin, 1995; Yuan & Ellis, 2003). This mechanism also allows teachers/researchers to elicit learners' answers during their task-based practices throughout various communication task types, such as the two-way information gap or jigsaw tasks (Chen, 2016a, 2018; Peterson, 2006; Smith, 2003).

Ortega and González-Lloret (2015) proposed a technology-mediated TBLT framework to operationalize task-based design in the virtual world to invigorate experiential learning that transcends the classroom. Three criteria are proposed for the dynamic duo (i.e. tasks and technology) to function: (1) real-life tasks need to be authentic in nature rather than camouflaged as exercise-based activities simply delivered to a digital platform; (2) teacher educators should consider wider applications of technology-mediated TBLT to language education and SLA; and (3) rigorous task-based design supported by relevant technology needs to follow a full task cycle, such as conducting needs analysis, selecting/sequencing tasks, and evaluating learners' task performance (González-Lloret & Ortega, 2014). Hence, meaningful real-life tasks that might seem cumbersome or challenging when carried out in a traditional class (e.g. checking into a hotel or taking multiple field trips in a target country) can be seamlessly configured in SL, thus mitigating issues related to travel arrangements and budget concerns (Canto *et al.*, 2014; Lee & Gerber, 2013; Peterson, 2012).

Given that language practices are developmental and complex in nature, a full-fledged virtual course extending over a longer time span was conducted in this study to capture the dynamic learning spectrum (Peterson, 2006, 2010a, 2010b). It not only enabled students to fully immerse themselves in the virtual community of practice, but also allowed the researcher to consistently observe and document the students' language practices for richer data. A task-based syllabus that addressed students' needs, promoted spontaneous interaction, and incorporated real-life scenarios into SL was adopted in this virtual course under the principles of TBLT design (Doughty & Long, 2003; González-Lloret & Ortega, 2014; Nunan, 2006; Skehan, 2003; see section 4.2, Data collection, for detail). The implementation of the TBLT syllabus that consistently documented task delivery and how learners' oral performance played out in SL further provides practical implications for language teachers interested in task-based instruction using 3D MUVEs (Chen, 2016b, 2018).

2.2 Pre-task planning

Ellis (1987) conceptualizes pre-task planning as an opportunity for L2 learners to plan and monitor their language output by restructuring their linguistic repertoire, which would help them ease into different levels of task demand (e.g. drafting a story outline before telling it orally). He further categorizes pre-task planning into rehearsal ("planning takes the form of an opportunity to perform the complete task once before performing it a second time") and strategic planning ("planning what content to express and what language to use but without opportunity to rehearse the complete task"; Ellis, 2009a: 474). Crookes (1989) also gives credit to its impact on fostering the learners' interlanguage processes because planned outputs can stretch learners' interlanguage system through planning and monitoring their language production. As L2 learners are generally expected to demonstrate their language proficiency against task performance, Skehan (1996) argues planning can be operationalized in "manipulable" task conditions in order to control

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

the cognitive load placed on L2 learners. This points to the importance of planning for learners to access, process, and organize their language resources before being tasked – in line with Ellis's (1987) conceptualization of pre-task planning.

Hence, the opportunity to plan before the tasks holds the potential to lower the learners' affective filter during pressed communication time as well as free up their cognitive capacity to attend to linguistic forms (Mehnert, 1998; Skehan & Foster, 2005). Those "hot spots" in L2 learners' interlanguage system that are not fully proceduralized will then have a better chance to be accessed and restructured under controlled task conditions for planning. In this sense, those difficult forms could be remedied and modified, leading to a more refined planned production (Foster & Skehan, 1996; Tajima, 2003; Yuan & Ellis, 2003). Although previous studies have argued that pre-task planning could optimize planned output in fluency, accuracy, and complexity, Ortega's (1999) studies on the effects of planning on interlanguage development reveals a mixed result – coinciding with Ellis's (2009a) comprehensive review. That is, a planned output is more conducive to fluent and syntactically complex quality than an unplanned output, but it is inconclusive in terms of accuracy. Despite this, Ortega (1999) argues

> . . . the theoretically interesting claim is not only that planning may lessen the cognitive load of a given task and free up attentional resources at the micro levels of speech production but also that it may foster during the planning phase a shift of conscious attention to formal aspects of the language needed to accomplish the task. (Ortega, 1999: 110)

We can hypothesize that planning allows for learners' attentional resources to be freed up for use and cognitive load can be lessened without time pressure. O'Malley and Chamot (1990) also stress that language learning is a complex cognitive process that takes both declarative knowledge (what we know) and procedural knowledge (what we know how to do) of the learner's interlanguage mechanism for learning to take place. It also requires practice to proceduralize the declarative knowledge (e.g. grammar rules, vocabulary) into a spontaneous "stage" for the purpose of language use (e.g. communication).

While pre-task planning appears pedagogically sound, there is still a paucity of studies on marrying SL and pre-task planning, thus deserving a closer observation in the current 3D MUVE and SLA literature. It would offer research and practical implications for SLA stakeholders to examine how pre-task planning as an instructional condition can impact learners' planned task outcomes in their oral output. Given the unique SL features that afford immersive simulation, real-time interaction, avatar-enabled tele/copresence, and multimodal communication, task-based design can be operationalized in 3D form to facilitate task delivery that transcends physical boundaries (González-Lloret, 2015; Jauregi, 2016; Ortega & González-Lloret, 2015; Peterson, 2016b; Wigham & Chanier, 2015). In response to Ortega's (1999) and Ellis's (2009a) mixed results, this study aims to investigate whether pre-task planning makes a difference in EFL learners' voice-based task performance in SL.

3. Research questions

The key question addressed in this study is, "*Does pre-task planning make a difference in EFL learners' oral performance as measured by complexity and accuracy in a task-based class conducted in SL?*" To address this, the quality and quantity of EFL learners' oral outputs during their task-based performance in the virtual course were measured at the levels of complexity and accuracy (Yuan & Ellis, 2003; see also section 4.3, Data analysis). In order to document their ongoing language practice in this virtual class, discourse samples selected from each student's oral task presentations were collected to provide empirical evidence on the effects of pre-task planning on learners' oral production in SL.

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

Figure 1. A snapshot of VIRTLANTIS island in SL

4. Methodology

4.1 Setting, participants, and ethics

VIRTLANTIS, a 3D island in SL, was selected as the research site for the virtual class, as it provides free language classes that attract SL residents to learn different foreign languages. Teacher volunteers are also present to offer classes to help students improve their target language proficiency. The major factor that distinguishes online teaching in VIRTLANTIS from teaching in real life is that it enables language teachers and learners to maximize real-life task-learning experience by using the salient SL features. For instance, learners in avatar form can build 3D objects in a *Sandbox, teleport/fly* to various spaces for social events, or simulate a real-life scenario (e.g. dining at a restaurant) by *rezzing* the *Holodeck* feature (i.e. creating or dragging a 3D object to the ground to make it appear; see Figure 1).

Snowball sampling (a non-random sampling procedure to encourage potential participants to spread the word to other like-minded participants) was employed via an invitation notecard sent to all VIRTLANTIS members regarding the nature of the task-based course and purpose of this study. Nine EFL learners who were keen on improving their English oral communication skills expressed interest in joining the class. They initially met with the researcher (also the teacher) in VIRTLANTIS for a one-on-one debriefing session to obtain their informed consent. They were informed that their oral production in each session would be audio-recorded for research purposes and their real-life identities would be kept intact, as their avatar names were not directly linked to personal names. They could also withdraw from the study at any time without obligations. Students were adult residents in SL (aged between 21 and 60) and came from diverse linguistic and cultural backgrounds (e.g. French, Spanish, Arabic, Indian, and Swedish). All were familiar with SL features and knew how to use voice/text chat to communicate with other avatars. Following the proficiency standards of the American Council on the Teaching of Foreign Languages (Swender, Conrad & Vicars, 2012), the researcher had the chance to assess each learner's English proficiency during their oral interview at the debriefing session (e.g. "*Why do you come to SL for practicing English?*") and their oral performance in pre-course task-based interaction. Their levels of language proficiency ranged from novice-high to intermediate-high, as assessed. Figure 2 presents a full version of each learner's demographic information, language learning background, English proficiency level, and intention to participate in this SL class (see Chen, 2016a).

4.2 Data collection

The 10-session virtual class (1.5 hours per session, twice weekly) incorporated real-life tasks that were meaningful and engaging to the EFL learners (Ellis, 2000; Nunan, 2006; Skehan, 2003).

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

Avatar name (initials)	UG	MB	PK	TR	NM	UL	IL	EC	BL
Gender	M	F	M	F	M	M	F	F	F
Age Range	31-40	31-40	51-60	21-30	18-20	21-30	21-30	21-30	41-50
Native language	French	Arabic	Spanish	Arabic	Gujarati	Persian	Thai	Swedish	Spanish
English proficiency	Inter-low	Inter-mid	Inter-mid	Inter-mid	Inter-low	Inter-high	Inter-mid	Inter-mid	Novice-high
First exposure to English	Middle high school	N/A	From doing business in real life	Since 2007	Since 2002	In college (since 2007)	Kindergarten	10 years old	High school
Language classes taken before SL	3 (mostly German)	none	none	none	2 (only sit-in)	none	1	>3	none
Reason to attend this virtual course	Serious about taking it as a good experience	Improve English; speak more confidently	Course info looks professional & helpful	Improve English	Improve oral/listening proficiency & presentation skill in English	Improve English communicat-ion skills in a unique way	Exchange experience and knowledge with people around the world; learn new technology	Improve English	SL offers many educational opportunities; fun and easy to interact with others by just sitting home
Why spend time in SL	N/A	Improve English; learn new things; interact with friendly people	Practice English with people around the world	Improve listening and speaking skills	Make friends and talk to people around the world; practice English	Just created the avatar for this course (a newbie)	Love learning new things in SL (e.g., science, language, culture)	Required in a university course	Enjoy building objects, exploring playing and learning in SL; knew SL since 2008

Figure 2. Demographic information of the participants

The task design was motivated by the pedagogical framework of technology-mediated TBLT (Doughty & Long, 2003; Ortega & González-Lloret, 2015), targeting authentic real-life tasks, potential implications for language learning, appropriate selection of technology, needs analysis, and task evaluation (González-Lloret & Ortega, 2014). To illustrate, a task-based syllabus was adapted from a pilot version with an EFL student group in SL who shared similar backgrounds. That is, they were also adult EFL students with culturally/linguistically diverse backgrounds, had no prior task-based learning experience before the SL class and came to VIRTLANTIS to practice English with others around the world (also see Chen, 2016a, 2018). Based on actual implementation, lesson materials and task designs were modified to strengthen the content validity. Additionally, tasks were selected in accordance with the needs analysis of the students based on their responses to a post-course survey and opinions in the interview of the pilot study to ensure face validity (Chen, 2016b). For instance, selected tasks not only enabled them to simulate real-life scenarios they found meaningful and engaging, but were also built upon their cultural repertoires in tandem with SL affordances to facilitate task delivery and experiential learning, such as teleporting to field trip sites or wearing cultural outfits in avatar form. The detailed task-based syllabus can also be viewed in the supplementary material (Chen, 2016b).

In order to examine the effect of pre-task planning on their oral outputs, students were given rehearsal time for task planning (Ellis, 2009a) at home before orally presenting their work in SL. As this virtual course was not obligatory and the class was not conducted in a lab, allocating specific time for pre-task planning was not feasible or controllable in this case. Hence, the students were allowed to have as much time as needed to rehearse at home to merit this ecological approach (Eckerth, 2008). For instance, they would research information on how to be a tour guide leading the class to different SL landmarks that simulated tourist spots in their country (e.g. Le Mont-Saint-Michel). Similarly, they demonstrated how to cook a cultural dish (e.g. paella), showcased their cultural costume worn by their avatar (e.g. flamenco costume), or introduced an artifact to the audience in a 3D museum gallery as a curator (see Figures 3 and 4 for illustrative task

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

Figure 3. A student working as a tour guide to showcase her Egyptian home culture

Figure 4. A student presenting an artifact to the class as a sculpture gallery curator

examples). This course also coincided with González-Lloret and Ortega's (2014) suggestion of designing a language program in a complete cycle to document the trajectories of learners' task performance and gather empirical evidence of the measured accuracy and complexity of their language outcomes in SL.

4.3 Data analysis

This study adopted Yuan and Ellis's (2003) analytical framework that used T-units to measure the quality and quantity of EFL learners' oral productions over time. A T-unit is the shortest unit of a sentence that can stand alone grammatically – "a main clause and related subordinate clauses and nonclausal structures embedded in it" (Hunt, 1970, as cited in Pica & Doughty, 1985: 119). A T-unit analysis has also been used extensively as a device to measure the syntactic complexity and accuracy of learners' speaking or writing samples in SLA (Pica & Doughty, 1985; Young, 1995; Yuan & Ellis, 2003). An ANOVA with repeated measures was performed and post hoc Bonferroni tests were run later to determine the locations of the significance through pairwise comparisons if the F scores were statistically significant ($p < .05$). Effect sizes are reported

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

Table 1. Measured variables for complexity and accuracy using T-units

Level	Measure	T-units (ratio)
Complexity	*Syntactic complexity*	The ratio of clauses to T-units.
	Syntactic variety	The total number of different grammatical verb forms used in the task.
	Type-token ratio	The number of different words divided by the total number of words in each segment of 40 words. Mean scores for all segments are added as the total and divided by the total number of segments (also known as Mean Segmental Type-Token Ratio).
Accuracy	*Error-free clauses*	The percentage of clauses that contain no errors.
	Error-free T-units	The percentage of grammatical T-units divided by grammatical and ungrammatical T-units.
	Correct verb forms	The percentage of accurate used verb forms divided by the total number of used verb forms.

to further examine the strength and magnitude of this study (Brown, 2008; Dörnyei, 2007). Both independent and dependent variables are described in the following sections.

4.3.1 Independent variable

Throughout the 10-session task-based course, six of the tasks required students to make oral presentations in front of the class, as previously explained. However, some of the students did not complete every task due to their real-life commitments. In order to ensure consistency of data analysis, three different points of time in which all students ($N = 9$) were present to perform all of those tasks were selected: *session 4* (show and tell one's cultural outfit in avatar or digital poster form), *session 8* (work as a gallery curator), *session 9* (work as a tour guide). Each session was denoted as T1 (session 4), T2 (session 8), and T3 (session 9) based on the time progression for each task.

4.3.2 Dependent variables

Since the students' oral presentations were monologic in nature and contained few elliptical utterances, T-units rather than C-units were analyzed, as rationalized by Yuan and Ellis (2003). An adopted measure scheme is presented in Table 1 (for detailed specifications of each measured variable, see Yuan & Ellis, 2003: 13–14).

In order to ensure coding consistency, an SLA specialist was recruited to ensure intercoder reliability. Before independent coding, the researcher conducted several training sessions where instructions for coding schemes were clearly explained to the second coder and sample data were also provided for her to practice. After questions and concerns raised by the second coder were resolved, they independently coded 30% of the whole data set. Intercoder reliability was then calculated using the formula of an intraclass correlation coefficient, which was set at a two-way mixed model (i.e. raters were fixed) and the absolute agreement type suggested by McGraw and Wong (1996). A high level of agreement was reached at .994 of average measure intraclass correlation coefficient ($p < .01$). The discrepancies in both codings were further compared and discussed before the rest of the data were coded. Figure 5 illustrates a vivid coding example of a student's oral presentation (T1: show and tell) with color coding and data quantification following each coding category of measured T-units (see Table 1).

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

Table 2. Differences in complexity throughout three progressional sessions

	M (*SD*)						Post hoc (Bonferroni)		
	T1	T2	T3	*F*	*p*	Par. eta^2	T1–T2	T2–T3	T1–T3
Syntactic complexity	1.27 (.16)	1.46 (.14)	1.35 (.13)	6.383	.009*	.444	.035*	.253	.333
Syntactic variety	5.89 (1.54)	5.67 (2.83)	8.67 (1.22)	12.430	.001**	.608	1.000	.011*	.003*
Lexical variety	.79 (.04)	.77 (.05)	.79 (.04)	.931	.415	.104	.492	1.000	1.000

Note. T1 = session 4; T2 = session 8; T3 = session 9; par. eta^2 (η_p^2) = partial eta squared (effect size).
*$p < .05$.
**$p < .001$.

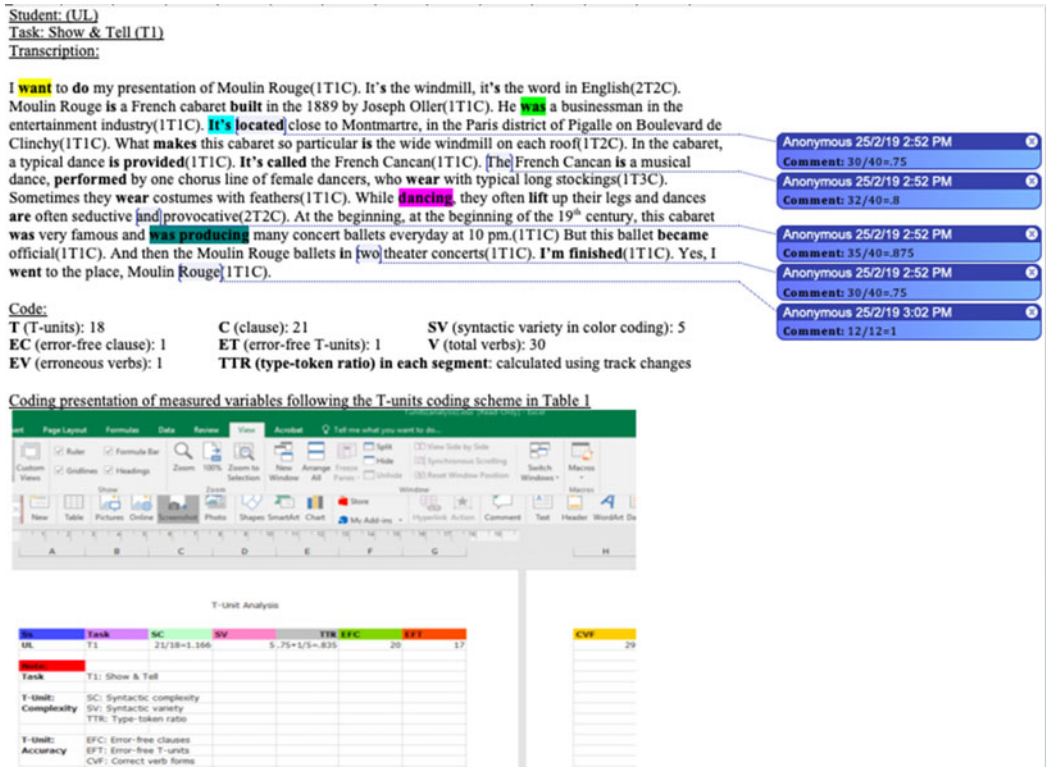

Figure 5. A coding example of the T-unit analysis

5. Results

5.1 Complexity

Table 1 shows three dependent variables tapping into the construct of complexity of the nine students' language output in the format of oral presentation: syntactic complexity, syntactic variety, and lexical variety (measured by Mean Segmental Type-Token Ratio). Results of the complexity level of students' oral production measured by each variable are presented in Table 2.

After performing a series of one-way repeated measures ANOVAs with Huynh–Feldt corrections (i.e. a more conservative procedure), the results showed that the overall differences in the means of the three dependent variables measuring complexity across the three sessions over time were statistically significant in syntactic complexity, $F(2, 16) = 6.38$, $p = .009$, and syntactic variety,

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

Table 3. Differences in accuracy across three progressional sessions

	M (*SD*)						Post hoc (Bonferroni)		
	T1	T2	T3	*F*	*p*	Par. eta^2	T1–T2	T2–T3	T1–T3
Correct clauses	32.67 (13.55)	29.89 (15.55)	62.11 (13.29)	14.157	.000**	.639	1.000	.005*	.003*
Correct T-units	22.89 (8.81)	20.00 (10.95)	46.67 (13.23)	13.615	.000**	.630	1.000	.008*	.006*
Correct verbs	51.33 (23.40)	45.11 (25.36)	107.89 (29.51)	20.922	.000**	.723	1.000	.003*	.001*

Note. T1 = session 2; T2 = session 8; T3 = session 10; par. eta^2 (η_p^2) = partial eta squared (effect size).
*$p < .05$.
**$p < .001$.

$F(2, 16) = 12.43$, $p = .001$, but not in lexical variety, $F(2, 16) = .93$, $p = .415$. The effect sizes also indicated that the magnitude of the difference was large, with 44.4% and 60.8% of the variance accounted for by syntactic complexity and syntactic variety respectively. In other words, EFL students throughout the course showed improvement in their grammatical complexity at the level of syntactic complexity and variety, but not as much in lexical sophistication measured by the variety of vocabulary use.

A further post hoc test using Bonferroni correction (to mitigate the statistical problem in multiple comparisons, such as ANOVAs) revealed that in the case of syntactic complexity, students in session 8 (T2) improved more than in session 4 (T1) at the statistically significant level ($p = .035$). However, the language output produced in the last session (T3) did not differ significantly from T1, although it was greater in mean ($M = 1.35 > M = 1.27$). The post hoc test on syntactic variety also showed a similar result in that students' latter output in the final session (T3) outperformed the two earlier sessions both at the statistically significant levels ($p = .011$ for the T2–T3 comparison, and $p = .003$ for the T1–T3 comparison). However, no difference was found when the first two sessions were compared, and the mean of the latter was slightly smaller than the former ($M = 5.67 < M = 5.89$). As the difference in lexical variety was not statistically significant, the post hoc test was not performed. Nevertheless, the means in the lexical variety of the three sessions were almost equal (.79), although slightly smaller (.77) in the second session (T2).

5.2 Accuracy

The accuracy of the students' oral production was measured based on the percentages of error-free clauses, error-free T-units, and correct use of verb forms. Table 3 summarizes the results.

Overall, statistically significant differences were found for all three dependent variables that measured the construct of accuracy in the students' oral production across the three sessions over time: $F(2, 16) = 14.157$, $p < .001$, for correct clauses; $F(2, 16) = 13.615$, $p < .001$, for correct T-units; $F(2, 16) = 20.922$, $p < .001$, for correct verbs. The effect sizes also indicated that the strength of the within-subject effect accounted for by each measured variable was considerably large (63.9% for correct clauses, 63.0% for correct T-units, and 72.3% for correct verbs). The positive findings indicated that the quality of the students' language output was greatly improved in terms of accuracy measured by error-free clauses, error-free T-units, and the correct use of verb forms.

Further post hoc tests pointed to the fact that statistically significant differences were located between the last session (T3) and the other two (T1, T2) in each measured variable, but no statistically significant differences were found between T1 and T2. In other words, students' oral production in terms of accuracy progressed greatly in the final sessions of the course. However, the mean of each variable in T2 was slightly lower than in T1, a finding that warrants further discussion below.

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

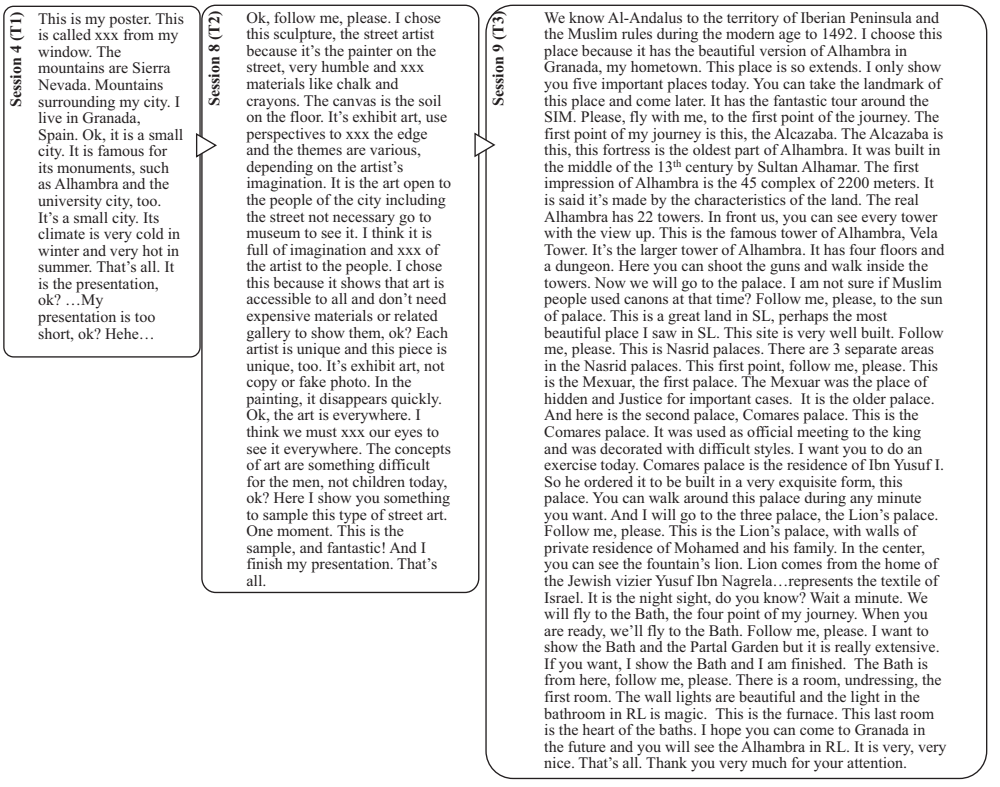

Figure 6. Excerpts of the developmental change in quality and quantity of student BL's oral output

5.3 Quality and quantity

The quality and quantity of EFL students' oral production did change over time throughout the task-based course. A vivid example is the language development of the beginner-level student (BL) throughout the course (see Figure 6). Not only did the quantity of her oral production in the final session increase more than the first two classes combined as measured by word counts (T3: 543 > T2: 196 > T1: 75), the language complexity also improved in relation to the number of different verbs used (T3: 7 > T1: 5 > T2: 4). More correct clauses (T3: 70 > T1: 22 > T2: 20), correct T-units (T3: 62 > T1: 18 > T2: 16) and correct verb forms (T3: 103 > T1: 32 > T2: 28) were also used.

6. Discussion

The results of learners' oral production were mixed. In terms of complexity of language output, the statistical results (as shown in Table 2) revealed that the EFL students showed marked improvement in the area of syntactic complexity and variety, but not in lexical variety. The results were also supported by the large effect sizes, with 44.4% and 60.8% of the variance accounted for in syntactic complexity and syntactic variety respectively. In terms of accuracy, the students' linguistic performance in the aspect of accuracy improved at all levels as the large effect sizes indicated 63.9%, 63.0%, and 72.3% of the variance were accounted for by error-free clauses, error-free T-units, and correct verb forms respectively. This finding was unexpected, as this task-oriented course was focused more on the process of task completion rather than on linguistic

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

elements (Long, 1990). Although "focus on form" could also be implemented in a task-based syllabus design (Ellis, 2009b; Long & Crookes, 1992), greater attention was paid to whether the students could finish their oral presentation followed by peer's and teacher's feedback on their overall performance given the time constraint in each session. It was also hypothesized that the sophistication of students' vocabulary use measured by their lexical variety would have outpaced their improvement in grammar since the "noticing" of learners' interlanguage processing was triggered mostly by lexical input, as previously discussed (e.g. Blake, 2000; Smith, 2003). Vocabulary acquisition, be it intentional or incidental, is also the building block of language development due to students' exposure to abundant lexical input from peers' oral presentations and teacher talk (Gass, 1999).

One possible explanation for this unexpected finding is the students' pre-task planning investment in each oral presentation. Unlike the unrehearsed task-based interaction in pre- and post-course interaction sessions (see Chen, 2016a, 2018), students were allowed to prepare their materials and carefully plan their linguistic performance at home before their oral presentations in SL. Regarding the complexity variable, the current study supports the claim that pre-task planning has a positive impact on grammatical complexity of the learners' language production (Crookes, 1989; Foster & Skehan, 1996; Ortega, 1999). In terms of grammatical complexity, the students showed improvement as measured by the ratio of clauses to T-units and syntactic variety, and by the total number of different verbs used. Mehnert (1998) also found that the more time allocated for pre-task planning (i.e. none vs. 10 minutes in her study), the better the complexity – this finding is supported by this study as students had more than 10 minutes to prepare at home before producing their oral output in SL. Skehan (1996) also argued that L2 learners usually have difficulties attending to both meaning and form simultaneously and need to compensate for allocating attentional resources to one aspect but not to both. Given the time to plan for modifying the linguistic aspects (e.g. syntactic and lexical), they have a better chance to improve the quality of their language output. As evidenced in this study, the EFL learners had more time to plan and revise their presentations, thus resulting in the use of more complex grammatical structures in oral presentations than those in previous studies.

Another surprising but positive finding is that the learners' linguistic performance improved on all accuracy levels, namely error-free clauses, error-free T-units, and correct verb forms (see Table 1). It is worth noting that this virtual course was developed following a task-based syllabus design instead of a grammar-based one. Grammatical errors made during the students' oral presentations were not corrected in order to keep the flow. The fact that the quality of their oral production excelled in accuracy across all measured levels was unexpected. This finding may support prior SLA research on pre-task planning in that allowing time for learners to plan before they are tasked will optimize the accuracy of their grammatical performance (see also Ellis, 1987; Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999; Skehan & Foster, 1997).

Even though EFL learners did produce more complex and accurate language output on the syntactic level in SL, the post hoc tests revealed mixed results of the locations of significance across the three progressional virtual sessions. Regarding the complexity measure, the last session (T3) outperformed the previous two (T1, T2) on syntactic variety, but the differences were not statistically significant on syntactic complexity and lexical variety, although the mean averages of T3 were slightly greater than the previous two. Interestingly, session 8 (T2) was hypothesized to outperform the previous session (T1), but it was not statistically significant on the measures of syntactic and lexical variety, albeit surpassing on syntactic complexity. Furthermore, the means in T2 were slightly lower than T1 on the last two measures on complexity. As for the accuracy measure, the locations of statistical significance performed by the post hoc tests quite consistently resided at the developmental trend where the last session (T3) outperformed the first two (i.e. $T3 > T1$; $T3 > T2$) across error-free clauses, error-free T-units, and correct verb forms. However, no statistically significant difference was found between T1 and T2 after the post

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

hoc tests. It turned out that the average means in T2 across the three measures were slightly lower than T1, which was unexpected and should have been otherwise.

The mixed results might have been confounded by the factor of interaction of time and task conditions despite the format of oral presentation being consistent in all three sessions. It would be hard to determine if the positive effect was due to the progression of time or the tasks in SL. However, when the task conditions in the three sessions were examined, it was found that students in session 8 (T2: work as a gallery curator) were allowed to work collaboratively in pairs due to the time constraint, whereas students in session 4 (T1: show and tell one's cultural outfit) and session 9 (T3: work as a tour guide) were mostly engaged in individual work as each student brought his/her own cultural expertise from their home country that might or might not be shared by others. In this case, the students' oral production in session 8 (T2) might be constrained by time (as each student only had limited time to finish his/her part); this could have lowered the overall ratio of T-units analysis in T2. This may also explain why T2 did not outperform T1 on syntactic and lexical variety, as the two levels were measured by the total numbers of different grammatical verb forms and type-token ratio.

7. Implications

7.1 The impact of pre-task planning on linguistic performance

In summary, a strong claim that the EFL students did develop better linguistic complexity over time cannot be made due to the confounding factor of the interaction of time and task conditions (i.e. students worked in pairs in T2, whereas in T1 and T3 sessions, they were mostly engaged in individual work). However, this study did reveal the statistically significant improvements on the EFL students' syntactic complexity and variety across all accuracy levels in oral performance. These positive results were owing to the pre-task planning effects on grammatical complexity of learners' oral production and grammatical accuracy (Ellis, 1987; Skehan & Foster, 1997, 2005; Tajima, 2003). This finding also supports the positive claim of prior research, namely allowing time for learners to plan before tasks will make a difference in the learners' linguistic performance measured by complexity and accuracy (Crookes, 1989; Ellis, 2009a; Mehnert, 1998; Ortega, 1999; Skehan, 1996; Yuan & Ellis, 2003). Therefore, it is suggested that pre-task planning be seeded in task-based instruction either in a classroom-based setting or a 3D MUVE in order to optimize the learners' language acquisition and production.

Above all, the quality and quantity of students' pre-task planning investment also hinges upon the level of task engagement and relevance perceived by the learners (see further discussion in section 7.3). As indicated previously, the unique features of SL afford learners to simulate real-world tasks that may be infeasible or cumbersome to conduct in a physical class, thus deepening their language immersion experiences in 3D form (Chen, 2016b, 2018; Cooke-Plagwitz, 2008, 2009; Peterson, 2016a, 2016b). These affordances make real-life task operationalization feasible and engaging, which in turn stimulates and sustains learners' task planning to better perform the tasks. Hence, this 3D MUVE approach offers a dynamic alternative to pre-task planning typically conducted in classroom or lab settings.

7.2 Learner investment propelled by positive peer pressure

An interesting phenomenon that was witnessed in the progression of the virtual course was "peer stimuli." Seeing well-prepared and well-articulated peer presentations over time motivated the learners to improve their output in order to be perceived as a "professional" by their peers. Student UL, for example, after seeing his colleagues' high-caliber presentations, requested that his oral presentation be postponed so that he could do more planning. The comments received on the weekly learning journals and in the post-course interview (see Chen, 2016b) also revealed that students spent a considerable amount of time researching online to strengthen the content of

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

their presentation. They would go to Wikipedia or other sites to search for more information about the topic in order to talk like a professional in front of the class. Consequently, they were also exposed to rich syntactic and lexical input from online materials. It was noted that the work and time that the students put into the pre-task planning also translated into richer content and more complex sentence structures and sophisticated vocabulary – although the latter did not factor as much as the former in the oral production measured by overall quality.

7.3 Task-based instruction in SL: Authentic, cultural, and simulated

Operationalized under TBLT, this 10-session virtual course was aimed at building connections between authentic tasks and students' experiential learning that "contrasts with a 'transmission' approach to education in which the learner acquires knowledge passively from the teacher" (Nunan, 2006: 12). Given the multicultural/lingual nature of this class, EFL students also brought their cultural, linguistic, and SL expertise to the class. Thus, tasks were purposefully designed to capitalize on their funds of knowledge and cultural repertoires that enabled simulation of real-life tasks and to promote learning by doing (Doughty & Long, 2003). Taking on their role as a "culture ambassador" also motivated them to promote their home culture and invest more time and effort into enhancing the quality of their presentation, by making it more accurate and professional. Increased engagement and commitment in the tasks assigned to them helped them move from the periphery toward the center of the virtual community of practices (Lave & Wenger, 1991) while having their investment of time and effort validated by the virtual community members (Norton, 2001).

Indeed, for these EFL students, doing real-life tasks that required "authentic" use of the target language for meaningful purposes was unfortunately confined to the classroom. The immersive and simulated nature of SL makes carrying out real-life tasks much easier, so much so that they could go on multiple virtual field trips within a few seconds, which would have been difficult to achieve in a conventional English class setting (Canto *et al*., 2014; Peterson, 2012). Simulated tasks not only deepened the EFL students' real-life task-learning experience, but also maximized their input acquisition of new knowledge and vocabulary, as represented in 3D scenes and objects (Chen, 2016a, 2018; Peterson, 2016a, 2016b). Immersing themselves in simulated real-life scenarios also removed their feeling of "being in a class," which heightened their motivation and engagement (Lee & Gerber, 2013; Wang & Burton, 2013). Hence, the ability to simulate real-life language immersion situations also enabled them to take ownership of discovering and constructing new knowledge (Dawley & Dede, 2014), promoting authentic target language use while providing rich exposure to language input enhanced by 3D, multimodal support (Cooke-Plagwitz, 2008, 2009). The findings of this study also echoed the positive claims in SL research that 3D MUVEs can transcend time and distance and promote experiential learning (Canto *et al*., 2014; Chen, 2016b). After all, the tasks should include activities that students will encounter in the real world, namely " . . . the hundred and one things people *do* in everyday life, at work, at play, and in between" (Long, 1985: 89).

8. Limitations and future directions

Analyzing the quality of students' oral production over time also indicated external factors that might have confounded the interaction of time and tasks (e.g. students might have also learned and used English outside the SL class), which, however, were not controllable in this research design. Unless it was a one-shot research design using tests in a lab setting, this confounding variable would still sneak into the data. In other words, it is hard to ascertain whether the positive effects were mostly due to the progression of time or to the tasks despite efforts made to keep tasks comparable throughout each session and the fact that statistically significant improvements were also evidenced as measured by complexity and accuracy. In order to avoid this pitfall in future

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

research, it is suggested that a counterbalanced measures approach be employed to cancel out the carryover (sequence) effects in future repeated measures design. In order to spread out the carryover effects (time*task) evenly over the measured conditions, it is also suggested that future studies use (1) random ordering of tasks for each student or (2) the counterbalanced measures design through *Latin Square* procedure for the sake of practicality. That is, the *tasks* will be rotated throughout each *course* session and students will be randomly divided into several *subgroups*.[1] Consequently, each student group will receive all the treatments (tasks) in exact form but in a different order for each group. By so doing, the effects of the confounding factor could be minimized since the variations have been spread out evenly over each condition.

It is also acknowledged that it would have been ideal to track how students had utilized their planning time, as pointed out by Ellis (2018). However, conducting a SL class where recruited students come from all over the world and live in different time zones is challenging, as evidenced in this case study. This factor also constrains the extent to which the researcher would have the capacity to track each student's pre-task planning in their own home as opposed to a lab setting where it is more controllable. Since this SL course was voluntary in nature, the participants were under no obligation to attend each session. While this aspect is ecologically and ethically sound, it also mirrors the reality that not every learner would consistently attend all the sessions for the purpose of data collection, unfortunately. Regardless, the recorded oral outputs in SL sessions provide solid evidence-based results showing the effects of pre-task planning on the quality and quantity of the students' oral performance, which is the essence of this study. That said, the aforementioned constraints, although inevitable and hence uncontrollable, prevented the researcher from designing a control group to consistently compare with the treatment group following the task-based approach. As no control group was utilized, it might be hard to ascertain the actual effects of pre-task planning found in this study with confidence, and other researchers and practitioners should use discretion when interpreting the findings vis-à-vis their own settings.

Plagiarism is another unforeseen factor that might have impacted the results of the students' oral productions. A couple of students inadvertently used parts of the web content they had researched in order to "enrich" their presentation – especially when the information sought was beyond their current cultural and world knowledge (e.g. the origin of Persian rugs). These students read their scripts word for word during their oral presentations, as opposed to spontaneous speaking and using notes as talking points. The issue of plagiarism was not evident until the researcher started to transcribe the data. Although the copied contents were removed from data analysis and only the students' own words were analyzed, the issue of plagiarism should have been raised with the students and taken into consideration in the research design.

9. Conclusion

This study epitomizes the positive impact of incorporating theoretically sound and pedagogically oriented task design on EFL learners' oral productions in SL. When tasks are related to the real world and tapping into learners' cultural, linguistic, and world knowledge, learners will become more engaged, motivated, and willing to tackle the demands of the task. Worth noting are the learning initiatives taken by the EFL students in this non-credit-bearing and obligation-free course. Since participants could see the connection between the SL tasks and their home cultures, they exploited their cultural resources to help them perform culture-embedded tasks in English that might be cognitively and linguistically challenging. Thus, authentic cultural tasks triggered improved language outputs and stimulated cognitive processes that drew upon students' background knowledge surrounding their home culture and the world (Duff, 1986). Given the simulated immersion and multimodal features in SL, real-life task simulation was not only feasible for the EFL students, but also enhanced experiential learning and deepened knowledge

[1]For the sake of data coding, tasks were denoted as T1, T2, T3; course sessions, C1, C2, C3; subgroups, S1, S2, S3.

https://doi.org/10.1017/S0958344020000026 Published online by Cambridge University Press

construction and input acquisition (Chun *et al*., 2016; Dawley & Dede, 2014; Peterson, 2016a, 2016b; Sadler & Dooly, 2013).

Finally, the development of EFL learners' language output measured by complexity and accuracy using T-units discourse analysis is still a relatively new avenue in 3D MUVEs. Hence, it is worth exploring the extent to which the quality and quantity of learners' task-based language output can progress over time in 3D MUVEs. Empirical evidence and discourse gathered from this study has made a case in this regard.

Supplementary material. To view supplementary materials referred to in this article, please visit https://doi.org/10.1017/S0958344020000026

Ethical statement. Please refer to section 4.1, Setting, participants, and ethics, in this article.

References

Blake, R. (2000) Computer mediated communication: A window on L2 Spanish interlanguage. *Language Learning & Technology*, 4(1): 111–125.

Brown, J. D. (2008) Effect size and eta squared. *JALT Testing & Evaluation SIG Newsletter*, 12(2): 38e43.

Canto, S., de Graaff, R. & Jauregi, K. (2014) Collaborative tasks for negotiation of intercultural meaning in virtual worlds and video-web communication. In González-Lloret, M. & Ortega, L. (eds.), *Technology-mediated TBLT: Researching technology and tasks*. Amsterdam: John Benjamins, 183–212. https://doi.org/10.1075/tblt.6.07can

Chen, J. C. (2016a) EFL learners' strategy use during task-based interaction in Second Life. *Australasian Journal of Educational Technology*, 32(3): 1–17. https://doi.org/10.14742/ajet.2306

Chen, J. C. (2016b) The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. *Computers & Education*, 102: 152–171. https://doi.org/10.1016/j.compedu.2016.08.004

Chen, J. C. (2018) The interplay of tasks, strategies and negotiations in Second Life. *Computer Assisted Language Learning*, 31(8): 960–986. https://doi.org/10.1080/09588221.2018.1466810

Chun, D., Kern, R. & Smith, B. (2016) Technology in language use, language teaching, and language learning. *The Modern Language Journal*, 100(S1): 64–80. https://doi.org/10.1111/modl.12302

Cooke-Plagwitz, J. (2008) New directions in CALL: An objective introduction to Second Life. *CALICO Journal*, 25(3): 547–557. https://doi.org/10.1558/cj.v25i3.547-557

Cooke-Plagwitz, J. (2009) A new language for the Net Generation: Why Second Life works for the Net Generation. In Oxford, R. & Oxford, J. (eds.), *Second language teaching and learning in the Net Generation*. Honolulu: National Foreign Language Resource Center, University of Hawai'i at Mānoa, 173–180.

Crookes, G. (1989) Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11(4): 367–383. https://doi.org/10.1017/s0272263100008391

Dawley, L. & Dede, C. (2014) Situated learning in virtual worlds and immersive simulations. In Spector, J. M., Merrill, M. D., Elen, J. & Bishop, M. J. (eds.), *Handbook of research on educational communications and technology* (4th ed.). New York: Springer, 723–734. https://doi.org/10.1007/978-1-4614-3185-5_58

Dörnyei, Z. (2007) *Research methods in applied linguistics*. New York: Oxford University Press.

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

Duff, P. (1986) Another look at interlanguage talk: Taking task to task. In Day, R. R. (ed.), *Talking to learn: Conversation in second language acquisition*. Rowley: Newbury House, 147–181.

Eckerth, J. (2008) Investigating consciousness-raising tasks: Pedagogically targeted and non-targeted learning gains. *International Journal of Applied Linguistics*, 18(2): 119–145. https://doi.org/10.1111/j.1473-4192.2008.00188.x

Ellis, R. (1987) Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition*, 9(1): 1–19. https://doi.org/10.1017/s0272263100006483

Ellis, R. (2000) Task-based research and language pedagogy. *Language Teaching Research*, 4(3): 193–220. https://doi.org/10.1177/136216880000400302

Ellis, R. (2009a) The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30(4): 474–509. https://doi.org/10.1093/applin/amp042

Ellis, R. (2009b) Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3): 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Ellis, R. (2018) *Reflections on task-based language teaching*. Bristol: Multilingual Matters. https://doi.org/10.21832/9781788920148

Foster, P. & Skehan, P. (1996) The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3): 299–323. https://doi.org/10.1017/s0272263100015047

Gass, S. (1999) Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21(2): 319–333. https://doi.org/10.1017/S0272263199002090

González-Lloret, M. (2015) *A practical guide to integrating technology into task-based language teaching*. Washington, DC: Georgetown University Press.

González-Lloret, M. & Ortega, L. (2014) *Technology-mediated TBLT: Researching technology and tasks*. Amsterdam: John Benjamins. https://doi.org/10.1075/tblt.6

Jauregi, K. (2016) Researching telecollaboration processes in foreign language education: Challenges and achievements. In Carrió-Pastor, M. L. (ed.), *Technology implementation in second language teaching and translation studies*. Singapore: Springer, 155–178. https://doi.org/10.1007/978-981-10-0572-5_9

Lave, J. & Wenger, E. (1991) *Situated learning: Legitimate peripheral participation*. Cambridge: University of Cambridge. https://doi.org/10.1017/cbo9780511815355

Lee, Y. J. J. & Gerber, H. (2013) It's a WoW World: Second language acquisition and massively multiplayer online gaming. *Multimedia-Assisted Language Learning*, 16(2): 53–70. https://doi.org/10.15702/mall.2013.16.2.53

Linden Lab. (2013) Infographic: 10 years of second life [Press release]. https://www.lindenlab.com/releases/infographic-10-years-of-second-life

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

Long, M. H. (1990) Task, groups, and task-group interactions. In Anivan, S. (ed.), *Language teaching methodology for the nineties (Anthology Series 24)*. Singapore: SEAMEO Regional Language Centre, 1–50.

Long, M. H. & Crookes, G. (1992) Three approaches to task-based syllabus design. *TESOL Quarterly*, 26(1): 27–56. https://doi.org/10.2307/3587368

McGraw, K. O. & Wong, S. P. (1996) Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1: 30–46. https://doi.org/10.1037/1082-989X.1.1.30

Mehnert, U. (1998) The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20(1): 83–108. https://doi.org/10.1017/s0272263198001041

Norton, B. (2001) Non-participation, imagined communities and the language classroom. In Breen, M. P. (ed.), *Learner contributions to language learning: New directions in research*. Harlow: Pearson Education, 159–171.

Nunan, D. (2006) Task-based language teaching in the Asia context: Defining 'task'. *Asian EFL Journal*, 8(3): 12–18.

O'Malley, J. M. & Chamot, A. U. (1990) *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139524490

Ortega (1999) Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21(1): 109–148. https://doi.org/10.1017/s0272263199001047

Ortega, L. & González-Lloret, M. (2015) Staking out the territory of technology-mediated TBLT. In Bygate, M. (ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference*. Amsterdam: John Benjamins, 59–86. https://doi.org/10.1075/tblt.8.03ort

Peterson, M. (2006) Learner interaction management in an avatar and chat-based virtual world. *Computer Assisted Language Learning*, 19(1): 79–103. https://doi.org/10.1080/09588220600804087

Peterson, M. (2010a) Learner participation patterns and strategy use in Second Life: An exploratory case study. *ReCALL*, 22(3): 273–292. https://doi.org/10.1017/S0958344010000169

Peterson, M. (2010b) Massively multiplayer online role-playing games as arenas for second language learning. *Computer Assisted Language Learning*, 23(5): 429–439. https://doi.org/10.1080/09588221.2010.520673

Peterson, M. (2012) EFL learner collaborative interaction in Second Life. *ReCALL*, 24(1): 20–39. https://doi.org/10.1017/s0958344011000279

Peterson, M. (2016a) The use of massively multiplayer online role-playing games in CALL: An analysis of research. *Computer Assisted Language Learning*, 29(7): 1181–1194. https://doi.org/10.1080/09588221.2016.1197949

Peterson, M. (2016b) Virtual worlds and language learning: An analysis of research. In Farr, F. & Murray, L. (eds.), *The Routledge handbook of language learning and technology*. New York: Routledge, 308–319.

Pica, T. & Doughty, C. J. (1985) Input and interaction in the communicative language classroom: A comparison of teacher-fronted and group activities. In Gass, S. M. & Madden, C. G. (eds.), *Input in second language acquisition*. Rowley: Newbury House Publishers, 115–132.

Prensky, M. (2005a) "Engage me or enrage me": What today's learners demand. *EDUCAUSE Review*, 40(5): 61–64.

Prensky, M. (2005b) Listen to the natives. *Educational Leadership*, 63(4): 8–13.

Puentedura, R. R. (2006) Transformation, technology, and education [Blog post]. http://hippasus.com/resources/tte

Sadler, R. (2012) Virtual worlds: An overview and pedagogical examination. *Bellaterra Journal of Teaching & Learning Language & Literature*, 5(1): 1–22. https://doi.org/10.5565/rev/jtl3.456

Sadler, R. & Dooly, M. (2013) Language learning in virtual worlds: Research and practice. In Thomas M., Reinders H. & Warschauer M. (eds.), *Contemporary computer-assisted language learning*. London, UK: Bloomsbury, 159–182.

Skehan, P. (1996) A framework for the implementation of task-based instruction. *Applied Linguistics*, 17(1): 38–62. https://doi.org/10.1093/applin/17.1.38

Skehan, P. (2003) Task-based instruction. *Language Teaching*, 36: 1–14. https://doi.org/10.1017/S026144480200188X
Skehan, P. & Foster, P. (1997) Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1(3): 185–211. https://doi.org/10.1177/136216889700100302
Skehan, P. & Foster, P. (2005) Strategic and on-line planning: The influence of surprise information and task time on second language performance. In Ellis, R. (ed.), *Planning and task performance in a second language*. Amsterdam: John Benjamins, 193–216. https://doi.org/10.1075/lllt.11.12ske
Smith, B. (2003) Computer-mediated negotiated interaction: An expanded model. *The Modern Language Journal*, 87(1): 38–57. https://doi.org/10.1111/1540-4781.00177
Swain, M. & Lapkin, S. (1995) Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3): 371–391. https://doi.org/10.1093/applin/16.3.371
Swender, E., Conrad, D. & Vicars, R. (2012) *ACTFL proficiency guidelines 2012*. Alexandria, VA: American Council for the Teaching of Foreign Languages.
Tajima, M. (2003) The effects of planning on oral performance of Japanese as a foreign language. Purdue University, unpublished PhD.
Wang, F. & Burton, J. K. (2013) Second Life in education: A review of publications from its launch to 2011. *British Journal of Educational Technology*, 44(3): 357–371. https://doi.org/10.1111/j.1467-8535.2012.01334.x
Wigham, C. R. & Chanier, T. (2015) Interactions between text chat and audio modalities for L2 communication and feedback in the synthetic world Second Life. *Computer Assisted Language Learning*, 28(3): 260–283. https://doi.org/10.1080/09588221.2013.851702
Young, R. (1995) Conversational styles in language proficiency interviews. *Language Learning*, 45(1): 3–42. https://doi.org/10.1111/j.1467-1770.1995.tb00961.x
Yuan, F. & Ellis, R. (2003) The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1): 1–27. https://doi.org/10.1093/applin/24.1.1

About the author

Julian ChengChiang Chen is a senior lecturer in the Master of Arts (Applied Linguistics) course at Curtin University, Australia. His research interests include technology-enhanced language learning, task-based language teaching, and action research. His work has appeared in flagship journals, such as *Computer Assisted Language Learning*, *Computers & Education*, *System*, *TESOL Quarterly* and *The Modern Language Journal*.

Author ORCiD. Julian ChengChiang Chen, https://orcid.org/0000-0001-7788-0462